Crosslinguistic differences in the online planning of Relative Clause production in Japanese and Spanish speakers: An analysis of language production through Eyetracking methodology.

Laura Rodrigo Cristóbal

Directores:
José Manuel Igoa González (Univ. Autónoma de Madrid)
Hiromu Sakai (Waseda University)

Tesis Doctoral
Cognición y Trastornos
Facultad de Psicología
Universidad Autónoma de Madrid
2017

Acknowledgements

There are so many people that have helped me along this process I don't even know where to start from. First of all, I want to thank my directors, José Manuel Igoa and Hiromu Sakai, for their unconditional support and diligent supervision. For offering me invaluable opportunities to growth and learn and for helping me whenever I was lost in this path, for their motivation and guidance. It has been a pleasure to work with them through all these years.

I also want to thank to the professors and researchers in the Department of Japanese Education and in the Department of Psychology in Hiroshima University, for their continuous interest and wise comments have helped me develop this work. In special I want to thank Prof. Yukiko Hatasa and Prof. Kazumitsu Chujo, for invaluable feedback and dedication.

To Yingyi Luo, a person whom I admire, for her support and advice, for being a continuous source of knowledge and inspiration. To Manami Sato and Takuya Kubo, for introducing me to the technology of Eye-tracking, and letting me learn with them.

To Prof. Matatoshi Koizumi, for allowing me to participate in the OS-languages project, a project that has make me growth both as a person and as a researcher. To all the people I have been able to collaborate in this project, in special to Michihiro Tanaka and Hajime Ono.

My most sincere gratitude as well to Waseda University for accepting me and make me feel I had always been part of the Department. In special, I want to thank Yasuaki Shinohara, Hinako Masuda and Rinus Verdonschot.

To all the people I have met in the Department of Japanese Education at Hiroshima University: Akiko Cho, Yin Shuai, Yunzhu Wang, Ryoko Kuroda, Ai Matsubara, Zeinab Shekerabi and many others I'll have troubles finding their English names! You know who you are. Thank you so much for being there, thank you for the long hours of research and life discussions.

Additionally, I want to thank the Department of Japanese of the EOI in Madrid, and in concrete, to Isabel San Gabino, for opening the doors of the school for me to, not only recruit my participants, but also run my experiments. Thank you very much to the CEAO department of Asian Studies at UAM as well, for letting me go into their classes to explain my project and recruit participants.

Finally, but not less important, I want to thank all the participants who have being part of this project, thank you not only for participating, but also for sharing with friends in my quest for participants, thank you for actively helping me in the battle: without you none of this would have been possible.

And of course, to all my dearest friends, both in Japan and in Spain, for being there, and for their unconditional support.

Abstract

Language production involves different steps in order to transform non-linguistic messages to speech. They include establishing structural (syntactic) relations among sentence constituents, and selecting the appropriate lexical items to convey the intended message. However, the precise way structural relations and lexical information are computed along the sentence formulation process is not clear. This work explores how the non-linguistic message undergoes linguistic encoding (i.e. what information is used first, and to what extent processes overlap), and how structural and lexical information intertwine (if at all) during the time-course of sentence formulation. We analyze the production planning of Relative Clauses (hereafter, RCs), in Spanish (head-initial language) and Japanese (head-final language) by monolingual speakers, and in Japanese by Spanish-Japanese late bilinguals, by means of the eye-tracking method while participants described colored pictures. Variables tested were the animacy of the agent and the patient of the event, and RC-type, with either the agent or the patient as head noun (HN). Moreover, we compared utterances with active voice and with passive voice. By using RCs in a head-final language, compared to a head-initial one, we were able to explore sentence planning in a structure where the syntactic most dominant element (the HN) is not the first element. This comparison allows teasing apart the tenets of hierarchical incrementality (prioritizing structural relations over lexical items) with those of linear incrementality (lexical items being selected first). Moreover, by comparing RCs with the patient as HN in active and passive voice in Japanese and Spanish we were able to test the effects of grammatical role assignment on the time-course of speech planning, when word order is constant. Finally, manipulating animacy allows us to explore the role of conceptually salient elements along the process. In turn, bilingual speakers have to adapt to a language in which the basic word order is reversed. They may exhibit different strategies of sentence planning that those shown by monolingual speakers or, alternatively, behave in a similar way to either group of monolinguals, showing only quantitative differences in the way sentences are planned.

Results with monolingual speakers showed Japanese speakers focus extensively on the HN before directing their gazes to the element they are going to utter first, suggesting a speech planning process closer to hierarchical incremental accounts, which prioritize structural relations. Spanish monolinguals showed a pattern in which both structural and linear information appears to be more related from the beginning. The comparison between active and passive sentences showed a late effect of voice in both languages, with more effortful integration of grammatical roles in RCs with the patient as HN uttered in active voice. This effect suggests that grammatical roles are not fully assigned beforehand in either language, and assignment is not fulfilled until lexical retrieval has started. Bilingual speakers exhibited the same pattern as Japanese monolinguals: a planning process starting with the HN before retrieving lexical items, and with grammatical roles being fully assigned at late stages. Implications for a flexible system that gives priority to structural relations but allows flexibility between and within speakers are discussed.

Resumen

Hablar conlleva una serie de pasos que median entre el mensaje no-lingüístico y el habla articulada. Estos incluyen establecer relaciones estructurales (sintácticas) entre constituyentes, y seleccionar el léxico apropiado para expresar el mensaje. Sin embargo, no está clara la forma en la que las relaciones estructurales y la información léxica se computan a lo largo del proceso de planificación. Este trabajo explora cómo tiene lugar la codificación lingüística (qué información se usa en primer lugar, y en qué medida los procesos se solapan), y cómo la informacion estructural y la información léxica se entrelazan (si lo hacen) durante el desarrollo temporal de la formulación del habla. Analizamos la planificación de las Oraciones de Relativo (en adelante, ORs), en español (lengua de núcleo inicial) y en japonés (lengua de núcleo final) en monolingües, y en japonés en bilingües español-japonés, mientras rastreamos los movimientos oculares de los participantes al describir dibujos. Las variables analizadas fueron la animacidad del agente y del paciente y el tipo de OR, con el agente o el paciente como núcleo. Adicionalmente, comparamos oraciones expresadas en voz activa y pasiva. La comparación de ORs en una lengua de núcleo final y una de núcleo inicial nos permite explorar la planificación de la oración en una estructura en la que el elemento sintáctico más dominante (el núcleo) no es el primer elemento. De este modo, podemos diferenciar entre los principios de incrementalidad jerárquica (los cuales dan prioridad a las relaciones estructurales sobre los elementos léxicos) y los de incrementalidad lineal (con los elementos léxicos seleccionados en primer lugar). A su vez, la comparación de ORs con el paciente como núcleo en voz activa y pasiva en japonés y en español permite evaluar los efectos de la asignación de rol gramatical a lo largo del desarrollo de la planificación del habla cuando el orden de palabras es constante. Finalmente, con la manipulación de la animacidad exploramos el rol de los elementos conceptualmente salientes a lo largo del proceso. Por su parte, los hablantes bilingües se enfrentan a la tarea de adaptarse a una lengua cuyo orden de palabras básico es el opuesto. Pueden presentar estrategias de planificación lingüística diferentes a las de los monolingües o, alternativamente, pueden mostrar comportamientos similares a ambos grupos de monolingües, presentando sólo diferencias cuantitativas en la planificación de oraciones.

Los resultados con monolingües mostraron que los hablantes japoneses enfocan extensamente el núcleo antes de dirigir su mirada al elemento que será articulado primero, sugiriendo un proceso de planificación más cercano a la incrementalidad jerárquica, el cual da prioridad a las relaciones estructurales. Los monolingües españoles mostraron un patrón en el cual la información lineal y la estructural están más relacionadas desde el comienzo. La comparación entre oraciones activas y pasivas mostró un efecto tardío de la voz en ambas lenguas, lo que deja entrever un proceso de integración de los roles gramaticales más costoso en ORs con el paciente como núcleo en voz activa. Este efecto sugiere que los roles gramaticales no se asignan completamente a priori en ninguna de las dos lenguas, y que la asignación no se completa hasta que el acceso léxico ha empezado. Los bilingües exhibieron el mismo patrón que los monolingües japoneses: un proceso de planificación centrado en el núcleo antes de acceder al léxico, y con una asignación tardía de los roles gramaticales. Los resultados apuntan a un modelo flexible, que da prioridad a las relaciones estructurales, pero que permite flexibilidad entre hablantes y también dentro de los mismos hablantes.

Index of Contents

List of figures

List of tables

Introduction

Producing language is a task that takes a great deal of time in our daily lives. We are used to producing sentences, simple or complex, as a reply to a question, as a means to capture a potential listener's attention, and for many other reasons. This familiarity does not mean, however, that this task is not challenging for the speaker, as learners of a second language would readily admit.

When speakers want to produce a sentence, they have to select the proper vocabulary from a mental lexicon of more than 50000 words (Dijkstra, 2007), and put these words inside a certain frame that makes sense in the language they are using. This frame is constructed every time a sentence is produced, giving rise to the possibility of infinite different utterances. However, this herculean task has to take place quickly to allow efficient communication without continuous disfluencies; a requirement that native speakers continually fulfill without major problems. Incrementality in language production, the step-by-step formulation of linguistic expressions, lies behind this efficient process (Kempen & Hoenkamp, 1987; Levelt, 1989). Language production takes place along a series of steps, starting from the conceptualization of a pre-linguistic message to the final articulation of the well-formed utterance. Between both ends, linguistic encoding takes place, a process that, in turn, is divided between the creation of the structure that will constitute the skeleton of the speaker's speech and the selection of the proper words that will be placed in order (e.g. Levelt, 1989; Ferreira & Slevc, 2007; Ferreira, 2010).

Despite the importance of language production, little is known about it. This problem comes about mainly due to methodological difficulties. There is a wide variety of tasks and procedures that allow to explore language comprehension as it takes place (i.e. online) from different levels of analysis: for example, by measuring neural activity by means of ERP or fMRI, by using self-paced reading or click-detection tasks, or by analyzing eye-tracking measures, both during reading and in the visual-world paradigm. However, this is not the case with sentence production, where many of these techniques were not available to analyze language production in a non-invasive way as the participant is speaking. Fortunately, the development of eye-trackers that do not require fixed head for recording allowed to analyze speech planning as it unfolds. The last decade, and specially the last years, have witnessed a sharp increase in the interest towards speech production and the online planning that takes place before articulation.

In the studies presented in this thesis, we present an attempt to explore the mechanisms of language planning from an online perspective, by analyzing speakers' gazes to the different elements on a scene that the speaker has to describe.

Eye-tracking in sentence production

In order to measure sentence planning, most studies have used a visual world paradigm, in which speakers see a display with a picture on it, and have to describe the scene while eye movements are recorded (see Meyer, 2004, for a review). A tight relation between eye movements and speaking has been systematically reported (e.g. Meyer, Sleiderink & Levelt, 1998; Griffin & Bock, 2000). Speakers look at the items (objects or characters) they are going to produce consistently around 900 ms. before (in English) they name them. This span (known as the eye-voice span) is longer that the gazes devoted to recognizing objects, which can be as short as 170 ms. or to search for objects, which takes no longer than 300 ms. (Griffin, 2004).

Moreover, when there is more than one element to be produced, speakers consistently do not look at the element they are articulating at that moment, but they start looking at the second element in order to prepare it. Importantly, this has been found in easy tasks consisting of naming different objects (e.g. "The cup is next to the pen") (Meyer *et al*., 1998; Griffin, 2001), and in the production of transitive actions (e.g. "The turtle is squirting the mouse with water") (Griffin & Bock, 2000). In either of these cases, speakers fixate each of the elements in the same order of speech. These gazes, known as *name-related gazes* (Griffin, 2004), have been proposed to reflect the access to the lexical items and their phonological properties. Although these gazes only reflect the access to the lexical items, recent studies have proposed that there are gazes that are related to the preparation of the structure that will guide subsequent gazes to the elements in order. However, it is still not clear whether there is a preparation of the skeletal compounds of the utterance behorehand, and how this is reflected in gaze patterns. Griffin (2004) notes that these gazes for the structure might exist, but are difficult to grasp.

The main problem to differentiate gazes due to structure planning from those due to lexical retrieval rests on two reasons. First, the action –or the verb– is usually not a measurable entity to which speakers can direct their gazes (but see Hwang & Kaiser, 2014 for an attempt to measure gazes to verb regions), and it is not clear which elements speakers focus on when constructing a syntactic representation: the proposed hypothesis is that speakers focus to a similar extent on the various elements of the scene when they are preparing the structure (Norcliffe, Konopka, Brown & Levinson, 2015; Norcliffe & Konopka, 2015; Konopka & Kuchinsky, 2015; Griffin & Bock, 2000). However, gazes extensively directed to the agent have been found when preparing the verb (e.g. Norcliffe *et al*., 2015; Kubo, 2016). The second reason lies in the fact that the subject, which is the most dominant element of the sentence (to the same extent as the VP), appears at the beginning of it in most of the languages studied so far. The subject is an entity to which participants can direct their gazes, and so could be measured, but gazes directed to this element are confounded between structure construction and lexical access.

The present dissertation

Our study aims to disentangle structure-building processes from lexical retrieval of the elements in order. In order to do that, we focus on a structure in which the most syntactically dominant element is both a noun, an entity that will allow us to analyze gazes directed to it, and is not place at the beginning at the sentence, thus allowing to differentiate between name-related gazes and structure-related gazes. We chose to compare relative clauses (hereafter –RCs) in a head-initial and a head-final language: Spanish and Japanese, respectively.

In RCs, the syntactically most dominant element is the Head Noun (HN), that is, the element of which something is predicated, as can be seen in (1).

(1) The girl who ate an ice cream...

This element, in our example “girl”, is always a noun, and syntactically controls the whole relative clause, as this is a subordinate clause modifying that noun. In head initial languages, like English or Spanish, the first mentioned element and the most dominant one are again confounded. However, as it will be explained in more detail in Chapter 1, this is not the case for Japanese, a language that places the head noun after its modifying clause, as can be seen in (2).

(2) アイスクリームを食べた女の子…

Aisukuriimu-o tabeta onna-no-ko...

Ice cream-ACC ate girl...

In this dissertation, a comparison in RC planning between Spanish and Japanese will be presented. The focus is not only placed on analyzing two languages with different word orders but exactly same grammatical functions (as the comparison between (1) and (2) shows), but also on exploring the planning process underlying two structures with exactly the same word order but different grammatical functions. Again the aim is to analyze whether there is structural planning while name-related gazes have started, and whether both processes are discernible. This comparison is also possible with RCs, in Spanish and in Japanese; in this case, by looking at RCs with the patient as HN and with contrasting voice: active vs. passives. As can be seen in (3) and (4), in these two languages the word order is completely the same regardless of the voice that is selected, but neither the grammatical role of each of the elements involved in the action nor the verb morphology remains the same.

(3)

a. Spanish active:

El helado que comió la niña...

The ice cream that ate the girl...

“The ice cream that the girl ate...”

b. Spanish passive:

El helado que fue comido por la niña...

The ice cream that was eaten by the girl...

(4)

a. Japanese active:

女の子が食べたアイスクリーム…

Onna-no-ko-ga tabeta aisukuriimu...

Girl-NOM ate ice cream...

“The ice cream that the girl ate...”

b. Japanese passive:

女の子に食べられたアイスクリーム…

Onna-no-ko-ni taberareta aisukuriimu...

Girl-DAT was eaten ice cream...

“The ice cream that was eaten by the girl...”

One last factor that was controlled for along the studies presented in this dissertation was the animacy of agent and patient, a feature that contributes to the conceptual saliency of referents, as conceptual saliency modulates the voice of the final utterance, thus modifying grammatical role assignment (e.g. Bock & Warren, 1985; McDonald, Bock & Kelly, 1993; Prat-Sala & Branigan, 2000). We wanted to compare the effects of animacy on sentence planning.

In this dissertation we will present three studies that incorporate the three above mentioned variables: animacy, differences in the order of the HN, and voice contrast, as a means to explore different stages in sentence planning. Studies 1 and 2 were designed to address monolingual sentence planning, with Spanish monolingual speakers in the first case, and Japanese monolingual speakers in the second. Study 3, in contrast, was intended to analyze bilingual sentence planning, by analyzing RC planning by Spanish-Japanese late advanced bilinguals. The major aims of the three studies can be summarized as follows:

1. To explore how conceptually accessible elements influence sentence planning: do animate elements capture the speaker’s attention right from the beginning of the process of linguistic encoding due to their saliency, or is there an analysis of the whole scene regardless of the animacy of its components? Is there any difference between monolingual and bilingual speakers in the role given to conceptual saliency? In other words, our aim in this regard is to analyze the “starting point” in sentence planning (Bock, Irwin & Davidson, 2004): the information that is first delivered from the non-linguistic representation to the initial stage of linguistic encoding.
2. Related to the previous point, to explore the prioritized information in undertaking linguistic encoding. In other words, when undertaking linguistic encoding, are structural relations established first or is the planning of

individual lexical items that takes the lead? Is there any difference between a head initial and a head final language? How do bilingual speakers plan their speech in an L2 with a word order that is completely opposite to that in their L1?

3. To explore whether there is structural planning (grammatical role assignment) once name-related gazes have started, and to what extent it is differentiated from the retrieval of the lexical form of the words to be placed in order. That is, to explore to what extent relational (i.e., structural) and non-relational (i.e., lexical) information work in parallel along the sentence planning process.
4. To explore whether bilingual speakers show qualitatively different planning strategies than monolingual speakers: do they prioritize planning of the whole utterance over incremental planning? Is planning scope reduced for these speakers?

These questions refer to three different moments in the time-course of sentence planning, and will be addressed in order in the following chapters.

Structure of the dissertation

Along the six chapters of the dissertation, we will begin by reviewing the most relevant literature for our purposes, before introducing the empirical studies.

In Chapter 1, a comparison between RCs in Spanish and Japanese is presented. After a brief introduction of the studies that have addressed the asymmetry between subject RCs and object RCs, in both head initial and head final languages, a descriptive comparison of the features of RCs in Spanish and Japanese is presented. A special emphasis is given to the characteristics that will be central in this dissertation: difference in word order, paired with the similarities in syntactic construction between languages and, secondly, the stability of word order within languages in active and passive sentences. To discuss the linguistics of RCs is not the main aim of our work. These structures are a tool to analyze differences between and within languages in speech planning, allowing to disentangle easily confounded information in online speech planning: lexical and structural encodings. However, we consider it important to introduce the basic features of RC structures early, in order to frame the discussion that will follow.

After this brief introduction, in Chapter 2 a review of the literature in monolingual sentence production is presented. In the first place, we introduce the stages and processes involved in sentence production. After that, we explore a series of studies that have addressed the effects of conceptual accessibility on speech production, with special emphasis on animacy effects, the conceptual feature that was manipulated in our empirical studies. We also explore the importance of structural accessibility in language production and its repercussion, by reviewing studies that have made use of structural priming. Eye-tracking studies and planning scope studies have tried to analyze how incremental language production proceeds: what is the information that is used first, what part of the utterance is prepared before (and also after, in the case of eye-tracking studies) speech onset, and what are the upper and lower limits of advanced speech planning. Thus, studies coming from various methodological backgrounds will be evaluated.

Chapter 3 introduces the literature on bilingual sentence production. First, we present several theories about bilingual sentence production. These accounts differ in the extent to which they assume a relation between languages, in the role that proficiency plays on each model, and in the way structural relations are represented. However, they also differ in the level of analysis. We will briefly explore the main attempts to explain how bilingual speakers produce sentences. After that we will introduce studies that have explored structural accessibility in bilingual production: structural priming has been extensively used to explore the connections between languages at the structural and the word order levels. Here we will focus on the main findings. There follows a review of conceptual accessibility studies (note that the order of this review reverses the one followed in the review of monolingual research, mainly for expository reasons). Conceptual accessibility in bilingual speakers has been explored mainly from the standpoint of Second Language Acquisition studies, which focus on the role that animacy plays in aiding L2 learning. Very briefly, we will introduce the main findings from this field, focusing especially on Japanese. Research in the field of bilingual sentence production is much more scarce than in monolingual production, and to our knowledge, there are no attempts to explore bilingual sentence production by studying the issue of planning scope, and only one attempt so far using eye-tracking methodologies. We will address this study by Konopka and Forest (2016) before introducing the remaining questions in the field.

Chapters 4 and 5 introduce the empirical part of this dissertation. In Chapter 4 we describe studies 1 and 2. Study 1 was conducted with Spanish monolinguals while Study 2 was designed to explore RC production in Japanese monolinguals. Both studies are described in order. After them, a brief comparison of the results of both languages is carried out. Chapter 5 introduces the study run with bilingual speakers: native speakers of Spanish who had an advanced level of Japanese. The method and procedure of the three studies is the same: use of the visual-world paradigm to elicit RCs through a question directed to either the agent or the patient of the scene. There are slight variations between studies that will be presented in due course.

Chapter 6 presents the general discussion of the dissertation. We jointly address the results of the three studies: first, there is a discussion about the three temporal stages that were introduced when we described the aims of our current research. The chapter ends with a discussion of the implications and limitations of the three studies, along with some suggestions and plans for future research directions. Finally, this dissertation concludes with a brief section of Conclusions, that underscores the main findings of the present work.

Introducción

Hablar es una actividad que consume gran parte de nuestro tiempo en la vida diaria. Estamos acostumbrados a producir oraciones, simples o complejas, como respuesta a una pregunta, como una forma de captar la atención de un oyente potencial y por multitud de otras razones. Esta familiaridad, sin embargo, no significa que esta tarea venga desprovista de desafíos, como muchos aprendices de segundas lenguas habrán podido comprobar en primera persona.

Cuando un hablante quiere producir una oración, tiene que seleccionar el vocabulario adecuado de un lexicón mental de más de 50000 palabras (Dijkstra, 2009), y colocar estas palabras dentro de un marco que permita expresar un mensaje con sentido en la lengua que están usando. Este marco se construye cada vez que se produce una oración, lo que da lugar a la posibilidad de crear infinitas oraciones diferentes. Sin embargo, esta tarea hercúlea tiene que llevarse a cabo rápidamente para permitir una comunicación eficiente sin continuas difluencias, un requisito que los hablantes nativos de una lengua satisfacen continuamente sin mayores dificultades. Detrás de este eficiente proceso se encuentra el requisito de la incrementalidad en la producción del lenguaje, es decir, la necesidad de formular las expresiones lingüísticas paso a paso (Kempen y Hoenkamp, 1987; Levelt, 1989). La producción del lenguaje tiene lugar a lo largo de una serie de etapas, comenzando por la conceptualización de un mensaje pre-lingüístico hasta la articulación final de una oración bien formada. Entre ambos extremos tiene lugar la codificación lingüística, un proceso que, a su vez, está dividido en la creación de la estructura que constituirá el esqueleto del mensaje hablado y la selección de las palabras adecuadas para expresarlo que serán colocadas en un determinado orden (p.e. Levelt, 1989; Ferreira y Slevc, 2007; Ferreira, 2010).

A pesar de la importancia que tiene la producción del lenguaje, poco se sabe aún de este proceso. Este problema se origina, principalmente, en las dificultades metodológicas que entraña su estudio. Existe una gran variedad de tareas y procedimientos que permiten explorar la comprensión del lenguaje mientras tiene lugar (es decir, *online*, según la expresión inglesa) y desde diferentes niveles de análisis: por ejemplo, mediante la medición de la actividad neuronal a través de los potenciales relacionados con eventos (o ERP, en sus siglas en inglés) o la resonancia magnética funcional (o fMRI), mediante el uso de tareas de lectura auto-administrada o de detección de "clicks", o analizando los movimientos oculares tanto en la lectura como en el llamado "paradigma del mundo visual". Sin embargo, este no es el caso en la producción del lenguaje, donde no es posible usar muchas de estas técnicas para analizar la producción de una manera no invasiva mientras el participante está hablando. Afortunadamente, con el desarrollo de aparatos de registro de movimientos oculares que no requieren la inmovilización de la cabeza, se abre la posibilidad de analizar la planificación del habla al mismo tiempo que está teniendo lugar. La última década, y especialmente los últimos años, han presenciando un gran aumento en el interés por la producción del habla y la planificación *online* que tiene lugar antes de la articulación.

En los estudios que presentamos en esta tesis, ofrecemos un intento de explorar los mecanismos involucrados en la planificación del lenguaje desde una perspectiva *online*, mediante el análisis de las miradas a distintos elementos de una escena que los hablantes tienen que describir mediante una oración.

Registro de movimientos oculares en la producción del lenguaje

La mayoría de estudios enfocados a evaluar la planificación de oraciones han usado el procedimiento denominado "paradigma del mundo visual", en el cual los hablantes ven una pantalla con una imagen y tienen que describir dicha imagen mientras se graban sus movimientos oculares (véase Griffin, 2004, para una revisión). Sistemáticamente se ha observado una estrecha conexión entre los movimientos oculares y el habla (v.gr. Meyer, Sleiderink y Levelt, 1998; Griffin y Bock, 2000). Los hablantes dirigen su mirada a los elementos (objetos o personas) que van a producir 900 ms. aproximadamente antes de nombrarlos (en inglés). Este periodo (conocido como "eye-voice span" –periodo entre ojo y voz) es mayor que el periodo dedicado simplemente a reconocer objetos, que puede durar en torno a 170 ms., o el dedicado a localizar objetos, que no lleva más de 300 ms. (Griffin, 2004).

Por otra parte, cuando hay más de un elemento para producir, los hablantes sistemáticamente no miran al elemento que están articulando en ese momento, sino que comienzan a mirar al segundo elemento con el fin de prepararlo. Es de destacar que este patrón se ha observado tanto en tareas sencillas, consistentes en nombrar diferentes objetos (ej. "La taza está al lado del bolígrafo") (Meyer *et al.*, 1998;

Griffin, 2001), como en la producción de oraciones que describen eventos transitivos (ej. "La tortuga está mojando al ratón") (Griffin y Bock, 2000). En cualquiera de estos casos, los hablantes fijan cada uno de los elementos en el mismo orden en el que se producen. Estas miradas, conocidas como "miradas relacionadas con el nombre" (*name-related gazes*) (Griffin, 2004), parecen reflejar el acceso a las piezas léxicas y a sus propiedades fonológicas. Aunque estas fijaciones oculares reflejan únicamente el acceso al léxico, según propuestas más recientes también hay miradas destinadas a la preparación de la estructura, una actividad que guía las fijaciones posteriores a los elementos que se van a producir en el mismo orden en el que se articulan. Sin embargo, no está claro si se da una preparación del esqueleto de la oración a priori y, en tal caso, cómo se reflejaría esto en los patrones de fijaciones oculares. Griffin (2004) señala que las miradas destinadas a construir la estructura pueden existir, aunque son difíciles de detectar.

Dos razones distintas se encuentran detrás de la dificultad para distinguir las miradas debidas a la construcción de la estructura de aquellas destinadas a acceder al léxico. En primer lugar, la acción –o el verbo– normalmente no es una entidad medible a la que los hablantes puedan dirigir su mirada (sin embargo, véase Hwang y Kaiser, 2014, para un intento de medir las fijaciones dirigidas a las regiones de un dibujo relacionadas con el verbo), y no está claro cuál es el elemento en el que los hablantes se centran cuando están construyendo la representación sintáctica. La hipótesis propuesta es que los hablantes dirigen su mirada de forma equilibrada a varios elementos de la escena cuando están preparando la estructura (Norcliffe, Konopka, Brown y Levinson, 2015; Norcliffe y Konopka, 2015; Griffin y Bock, 2000). Sin embargo, también se han observado miradas dirigidas extensamente al agente cuando se está preparando el verbo (ej. Norcliffe *et al.*, 2015; Kubo, 2016). La segunda razón recae en el hecho de que el sujeto, que es el elemento sintácticamente dominante en la oración (al mismo nivel que el sintagma verbal), aparece al principio de la misma en la mayoría de las lenguas estudiadas hasta ahora. El sujeto es una entidad a la que los participantes pueden dirigir su mirada, y por tanto, puede ser evaluada, pero las miradas dirigidas a este elemento no permiten discriminar entre la construcción de la estructura y el acceso al léxico.

En el presente estudio tenemos como objetivo primordial distinguir los procesos de construcción de la estructura de los procesos debidos al acceso al léxico en el orden en que se van a producir las palabras. Para llevar a cabo esta tarea, nos centramos en una estructura en la que el elemento sintácticamente dominante es un nombre, es decir, una entidad que permite analizar las miradas dirigidas al objeto denotado por él, que además no aparece al principio de la oración, lo que nos permitirá diferenciar entre miradas relacionadas con el acceso al nombre y miradas relacionadas con la construcción de la estructura. Para ello, comparamos la producción de oraciones de relativo (en adelante, OR) en una lengua de núcleo inicial y otra de núcleo final: el español y el japonés, respectivamente.

En las ORs, el elemento sintácticamente dominante es el núcleo, es decir, el elemento del que se predica algo, como puede verse en (1).

(1) La niña que comió un helado...

Este elemento, "niña" en nuestro ejemplo, siempre es un nombre, y controla sintácticamente toda la cláusula de relativo, ya que ésta constituye una cláusula subordinada con función modificadora del nombre. En las lenguas de núcleo inicial, como el inglés o el español, el elemento que se menciona en primer lugar y el más dominante coinciden, por lo que estas dos características se pueden confundir. Sin embargo, como se explicará con más detalle a lo largo del Capítulo 1, este no es el caso del japonés, lengua que sitúa el núcleo después de la cláusula que lo modifica, como puede apreciarse en (2).

(2) アイスクリームを食べた女の子…
Aisukuriimu-o tabeta onna-no-ko...
Helado-ACC comió niña...

En esta tesis se llevará a cabo una comparación entre la planificación de ORs en español y en japonés. El foco no está centrado únicamente en el análisis comparado de dos lenguas con diferentes órdenes de palabras y exactamente las mismas funciones gramaticales (como muestra la comparación entre (1) y (2)), sino también en la exploración del proceso de planificación que subyace a dos estructuras que presentan exactamente el mismo orden de palabras, pero con funciones gramaticales diferentes. Con esta nueva comparación, nuestro objetivo es analizar si se produce una planificación estructural mientras tienen lugar las miradas relacionadas con el nombre y si ambos procesos son separables. Esto es posible mediante el análisis de ORs en español y en japonés, en este caso mediante el análisis de ORs que presentan el paciente como núcleo y se expresan en voz activa o pasiva. Como puede observarse en (3) y (4), en estas dos lenguas el orden se mantiene constante independientemente de la voz en la que la oración se exprese; sin embargo, este no es el caso del rol gramatical de cada uno de los elementos que participan en la acción, ni de la morfología verbal, los cuales varían.

(3)

a. Activa en español:

El helado que comió la niña...

b. Pasiva en español:

El helado que fue comido por la niña...

(4)

a. Activa en japonés:

女の子が食べたアイスクリーム…

Onna-no-ko-ga tabeta aisukuriimu...

Niña-NOM comió helado...

"El helado que comió la niña..."

b. Pasiva en japonés:

女の子に食べられたアイスクリーム…

Onna-no-ko-ni taberareta aisukuriimu...

Niña-DAT fue comido helado...

"El helado que fue comido por la niña..."

Un último factor que fue controlado a lo largo de los estudios que se presentan en este trabajo fue la "animacidad" de agente y paciente, una característica que contribuye a la "saliencia" conceptual de los referentes; numerosos estudios han hallado que la saliencia conceptual modula la voz de la oración emitida y, por tanto, modifica la asignación de los roles gramaticales (ej. Bock y Warren, 1985; McDonald, Bock y

Kelly, 1993; Prat-Sala y Branigan, 2000). Por ello, quisimos analizar los efectos de la animacidad en la planificación de oraciones.

En este trabajo presentaremos tres estudios que incorporan las tres variables mencionadas previamente: animacidad, diferencias en el orden de mención del núcleo y contraste de voz, como medida para explorar las diferentes etapas en la planificación de oraciones. Los estudios 1 y 2 tienen como objetivo explorar la planificación de oraciones en hablantes monolingües, centrándose para ello en monolingües de español, en el primer caso, y en monolingües de japonés, en el segundo. Por otro lado, el estudio 3 está dirigido a explorar la planificación de oraciones por parte de hablantes bilingües, y para ello analizamos la planificación de ORs por parte de bilingües tardíos español-japonés de nivel avanzado. Los objetivos principales de estos tres estudios se pueden resumir como sigue:

1. Explorar hasta qué punto y de qué modo los elementos conceptualmente más accesibles influyen en la planificación del habla: ¿los elementos animados capturan la atención del hablante desde el comienzo del proceso de codificación lingüística debido a su mayor saliencia o, por el contrario, hay un análisis previo de la escena completa, independientemente de la animacidad de cada uno de sus componentes? ¿Hay alguna diferencia entre hablantes monolingües y bilingües en el rol que desempeña la saliencia conceptual? En otras palabras, nuestro objetivo en este punto es analizar el "punto de inicio" en la planificación de oraciones (Bock, Irwin y Davidson, 2004): la información que es activada en primer lugar y remitida desde la representación no-lingüística a la fase inicial de codificación lingüística.
2. Relacionado con el punto anterior, explorar la información a la que se da prioridad al comienzo de la codificación lingüística. En otras palabras, al comienzo de la codificación lingüística, ¿las relaciones estructurales se establecen antes de que dé comienzo el proceso de selección léxica, o el proceso está dirigido por la planificación de piezas léxicas individuales? ¿Hay alguna diferencia entre una lengua de núcleo inicial y una de núcleo final? ¿Cómo planifican los hablantes bilingües el habla en una L2 en la cual el orden de palabras es completamente opuesto al de su L1?
3. Explorar si una de las operaciones implicadas en la planificación estructural, la asignación de roles gramaticales, tiene lugar una vez que las miradas

relacionadas con el nombre dan comienzo, y explorar, a su vez, en qué medida este proceso se diferencia del acceso al léxico en el orden de palabras tal como serán articuladas o se interrelaciona con él. En otras palabras, explorar en qué medida la información relacional (estructural) y la no-relacional (léxica) trabajan o no en paralelo a lo largo del proceso de planificación de oraciones.

4. Explorar si los hablantes bilingües muestran estrategias de planificación cualitativamente diferentes a las mostradas por los hablantes monolingües: ¿dan estos hablantes prioridad a la planificación de toda la oración por encima de la incrementalidad del proceso? ¿El alcance de la planificación se ve reducido en estos hablantes?

Estas cuestiones se refieren a tres momentos diferentes en el desarrollo temporal de la planificación del habla, y serán tratadas en orden en los capítulos que siguen.

Estructura de esta tesis

A lo largo de los seis capítulos que componen esta tesis, presentaremos inicialmente la literatura más relevante para nuestros objetivos, para a continuación exponer los estudios empíricos.

En el Capítulo 1, presentaremos una comparación entre las OR en español y en japonés. En primer lugar, se presentarán brevemente los principales estudios que han centrado sus esfuerzos en explicar la asimetría existente entre OR de sujeto y de objeto, tanto en lenguas de núcleo inicial como en lenguas de núcleo final. A continuación, se realizará una comparación meramente descriptiva de las principales características de las OR en español y en japonés. Se pondrá especial énfasis en aquellas características que resultan centrales en este trabajo: la diferencia en el orden de palabras, emparejada con las similitudes en las construcciones sintácticas entre lenguas; y, en segundo lugar, la estabilidad en el orden de palabras dentro de cada lengua en oraciones expresadas en voz activa y pasiva. La discusión de las características lingüísticas de las ORs no es el objetivo central de este trabajo. Estas estructuras se usan aquí como herramienta para analizar las diferencias en la planificación del habla, entre lenguas y también dentro de la misma lengua, ya que permiten discriminar tipos de información que normalmente se confunden o mezclan en la planificación *online* del habla: la codificación léxica y la estructural. Sin embargo, consideramos que es importante presentar las características básicas de las

OR al comienzo de este trabajo, para así poder enmarcar los capítulos y la discusión que seguirá esta tesis.

Tras esta breve presentación de la lingüística de las ORs, en el Capítulo 2 se realizará una revisión de la bibliografía más relevante en el campo de la producción de oraciones por parte de hablantes monolingües. En primer lugar, presentamos las fases y procesos que tienen lugar durante la producción de oraciones. Tras ello, exploramos una serie de estudios que han centrado sus esfuerzos en esclarecer los efectos de la accesibilidad conceptual en la producción del habla, con especial énfasis en los efectos de animacidad, la variable conceptual que se manipula en nuestros estudios empíricos. Asimismo, exploramos la importancia de la accesibilidad estructural en la producción del lenguaje y sus repercusiones, a través de una revisión de estudios que han hecho uso de la tarea de *priming* estructural. Por otro lado, los estudios basados en el registro de movimientos oculares y aquellos orientados a examinar el alcance de la planificación del habla han tratado de analizar cómo se desarrolla la producción incremental del lenguaje: qué información se usa antes, qué parte de la oración se prepara antes del comienzo del habla (y también después, en el caso de estudios de *eye-tracking*), y cuáles son los límites superiores e inferiores en la planificación del habla por adelantado (lo que en inglés se denomina *advanced speech planning*). Así, se evaluarán diversos estudios que hacen uso de estas metodologías.

El Capítulo 3 introduce la bibliografía más relevante en el campo de la producción de oraciones por parte de hablantes bilingües. En primer lugar, presentamos varias teorías acerca de la producción de oraciones en bilingües. Estas explicaciones difieren entre sí en la medida en que asumen distintas posturas sobre las relaciones entre lenguas, el rol que desempeña la competencia en cada uno de ellas y la manera en que se representan las relaciones estructurales. Asimismo, estos modelos también difieren en el nivel de análisis que adopta cada uno. Exploraremos brevemente los principales intentos de explicar cómo los hablantes bilingües producen oraciones. Tras ello, presentaremos estudios que han explorado la accesibilidad estructural en la producción en bilingües: la tarea de *priming* estructural ha sido ampliamente utilizada como forma de explorar las conexiones entre lenguas en los niveles estructural y de orden de palabras. Aquí nos centraremos en los principales resultados provenientes de este campo. A continuación, realizamos una revisión de estudios que han analizado los efectos de la accesibilidad conceptual en bilingües (nótese que el orden de esta revisión es el opuesto al presentado en el capítulo 2 con hablantes monolingües, principalmente para mayor claridad expositiva). La accesibilidad conceptual en hablantes bilingües ha sido explorada principalmente desde el punto de vista de los estudios de Adquisición de Segundas Lenguas, los cuales se centran en el rol de la animacidad como ayuda en el aprendizaje de la L2. De manera muy resumida, expondremos las principales conclusiones procedentes de este campo de investigación, centrándonos especialmente en el japonés. La investigación en el campo de la producción de oraciones en bilingües es mucho más reducida que en hablantes monolingües, y no ha habido intentos hasta la fecha, hasta donde alcanza nuestro conocimiento, de explorar la producción de oraciones en bilingües a través del estudio del alcance de la planificación, y tan sólo un intento hasta el momento de aplicar la metodología de registro de movimientos oculares para este propósito. Revisaremos este estudio, llevado a cabo por Konopka y Forest (2016), antes de exponer las cuestiones que quedan por resolver en este campo.

Los Capítulos 4 y 5 presentan la parte empírica de esta tesis. En el Capítulo 4 describimos los estudios 1 y 2. El estudio 1 fue llevado a cabo con monolingües de español, mientras que el estudio 2 estaba dirigido a explorar la producción de OR por parte de hablantes monolingües de japonés. Ambos estudios se describirán en este orden. Tras ellos, se llevará a cabo una breve comparación de los resultados en ambas lenguas. El Capítulo 5 presenta el estudio llevado a cabo en japonés (L2) con hablantes bilingües: nativos de español con un nivel avanzado de japonés. El método y el procedimiento son iguales en los tres estudios: todos ellos hicieron uso del paradigma del mundo visual para inducir la producción de ORs a través de una pregunta, referida bien al agente o bien al paciente de la escena. No obstante, hay pequeñas variaciones entre estudios que serán expuestas a lo largo de cada uno de ellos.

Finalmente, el Capítulo 6 presenta la discusión general de la tesis. En él, tratamos conjuntamente los resultados de los tres estudios: en primer lugar, la discusión se centra en los tres momentos temporales que se presentan en los objetivos de nuestra investigación (tanto al principio de esta introducción como, con mayor detalle, al comienzo de cada capítulo empírico). Tras ello, el capítulo concluye con una discusión de las implicaciones y limitaciones de los tres estudios, al tiempo que se presentan algunas sugerencias y planes para futuras investigaciones. La tesis concluye con un breve apartado de Conclusiones que destaca los principales hallazgos del presente trabajo.

Part 1.
Theoretical background

Chapter 1. Relative clauses in Spanish and Japanese: Main Characteristics

The sentence structure that is the focus of our study is relative clauses (hereafter, RCs), in two typologically different languages, namely Japanese and Spanish. RCs present interesting cross-linguistic differences that allow to explore, among other things, (1) the role that conceptual accessibility plays in sentence planning; and (2) the time-course of the mapping mechanisms between syntax and semantics and the processes regulating the flow of information between these components across different languages, and also in bilingual speakers in their two languages, and particularly in their L2.

RCs have been the focus of extensive attention in psycholinguistic research, both in language comprehension and speech production studies, since they are commonly used constructions, though they have a high syntactic complexity. In general, object RCs have been found to posit greater processing difficulty than subject RCs. However, there is still no consensus about why these differences arise, as we will explain in the next section.

Moreover, RCs have been the focus of research not only in order to explain intralinguistic variations, but also from a cross-linguistic perspective. Cross-linguistic variation in the formation of RCs does not only come in the order of constituents, but also in the linguistic foundations of the relationship between main and subordinate clauses. The analysis of the consequences that these differences might carry for parsers could provide valuable insights. RCs in Asian languages (Japanese, Korean and Chinese) seem to be governed by syntactic rules other than those governing Indo-european languages (although this point is still under debate) (e.g., Comrie, 2007; Ozeki & Shirai, 2007). For that reason, RCs have been an important field of research in these languages. On the other hand, despite cross-linguistic variation, word order within the RCs of a given language is usually quite stable, with few variations observed (MacDonald, 2013a). This aspect is especially important in a language like Japanese, which allows scrambling, or Spanish, that has relatively free word order (although it does not allow scrambling). This fact enables to carry out a more direct and reliable comparison across languages.

In the three studies we present in this dissertation, we explore the planning processes involved in RC production from a cross-linguistic perspective. We explore the differences stemming from the opposite word order in these constructions (head-initial: Spanish vs. head-final: Japanese) along with the differences concerning the dependency of RC processing on semantic cues (i.e., conceptual relations between elements) in each language. Hence, our main focus will be the comparison of object and subject RCs, and the manipulation of the animacy of their constituents.

Taking all this into account, it is necessary to begin with a brief explanation of how RCs are constructed in Spanish and Japanese, and what their main differences are, as well as a cursory review of the main differences found between subject and object RCs from a psycholinguistic perspective, before we present our study in full detail in the following chapters. Thus, this chapter will describe the main linguistic features of both Japanese and Spanish RCs.

Characteristics of Relative Clauses: A general view

Relative clauses, also known as adjectival clauses, function as modifiers of nominal expressions present in the matrix clause in which they are integrated (Fernández Lagunilla & Anula, 1995); in other words, RCs function as a modifier (adjective) of an NP of the sentence they are embedded in. This is the case not only for Spanish, but also for English (Biber, Johansson, Leech, Conrad & Finegan, 1999), Japanese (Iori, 2001), and most other languages.

All languages have relative clauses, but not all can relativize to the same extent. Thus, Keenan and Comrie (1977; Comrie & Keenan, 1979) postulated that all languages have the capacity to relativize at least subjects, but not all can relativize other kind of constituents in the sentence. The *Noun Phrase Accessibility Hypothesis* (hereafter, *NPAH*: Keenan & Comrie, 1977; Comrie & Keenan, 1979) establishes a hierarchy of accessibility of the different types of RCs that can be present in a given language. Through the study of over fifty languages, they proposed the distribution of RC constructions presented in (1), in which the least marked type and, consequently, the highest in the hierarchy are subject RCs (thus, universally shared by all languages), followed by direct object RCs, and so on:

(1)

Subject > Direct Object > Indirect object > Prepositional Object > Genitive > Object of Comparison

This hypothesis posits that the existence of a given type of RC structure in a language implies that all higher constructions in the hierarchy (to the left in the scheme presented in (1)) will be also present in that language. Some languages, like English, allow relativization at all levels, but this is not always the case: there is at least one language (from the set of 55 languages studied by the authors) that allows relativization only until each one of the intermediate levels of the hierarchy, thus including higher positions but excluding lower ones. Once the lowest position a language can relativize is known, all the RCs that that language allows can be predicted (e.g. a language that allows Indirect Object RCs will also relativize Direct Objects and Subjects, but not necessarily Prepositional Objects). Spanish also belongs to the group of languages that can relativize at all levels of the hierarchy (Keenan & Comrie, 1979) and Japanese can relativize to all levels except the lowest, that is, Object of Comparison (Iwasaki, 2002).

The NPAH thus offers a linguistic explanation of the observed asymmetry between subject RCs and object RCs. Subject RCs are (almost) universally easier to process than object RCs. This asymmetry has been observed in a great amount of languages, both head initial (English: e.g., King & Just, 1991; King & Kutas, 1995; Gordon, Hendrick, & Johnson, 2001; Traxler, Morris, & Seely, 2002; French: e.g., Cohen & Mehler, 1996; Dutch: e.g., Mak, Vonk, & Schriefers, 2002, 2006; German: e.g., Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kühn, 1995; and Spanish: Betancort, Carreiras, & Sturt, 2009) and head final languages (Chinese: Vasishth, Chen, Li & Guo, 2013; Jäger, Chen, Li, Lin & Vasishth, 2015; Korean: Kwon, Polinsky & Kluender, 2006; Kwon, Kluender, Kutas & Polinsky, 2013; Turkish: Kahraman, Sato, Ono & Sakai, 2010; Japanese: Ueno & Garnsey, 2008; Miyamoto & Nakamura, 2003; Ishizuka, Nakatani, & Gibson, 2003; Ishizuka, 2005).

Despite the fact that relative clauses are language universals (at least at the highest level of subject RCs), the way these are formed varies between languages, due to basic word orders, and other typological differences. This has led to a more complex panorama regarding the asymmetry between subject and object RCs in head-final languages. Studies with Chinese RCs show conflicting evidence, with some pointing towards the advantage of object RCs over subject RCs, due to word order, which is closer to canonical sentences (Chen, Ning, Bi & Dunlap, 2008; Hsiao & Gibson, 2003; Qiao, Shen & Forster, 2012; with context: Gibson & Wu, 2013; see Vasishth, 2015 for a review of different results in Chinese RCs). Moreover, in Japanese as well, there is evidence that, within an unambiguous context, object RCs are easier to process (Ishizuka, Nakatani & Gibson, 2006). Additionally, Basque is another head-final language (an ergative language) that has shown easier processing of object RCs (Carreiras, Duñabeitia, Vergara, de la Cruz-Pavía & Laka, 2010). However, in general terms, there is an overall consensus that object RCs yield greater costs for parsers in general. The reasons underlying this asymmetry have been widely explored.

From the NPAH perspective, subject RCs are higher in the hierarchy of accessibility and, thus, universally easier to grammaticize. However, the way in which this can be translated into ease of processing is not straightforward.

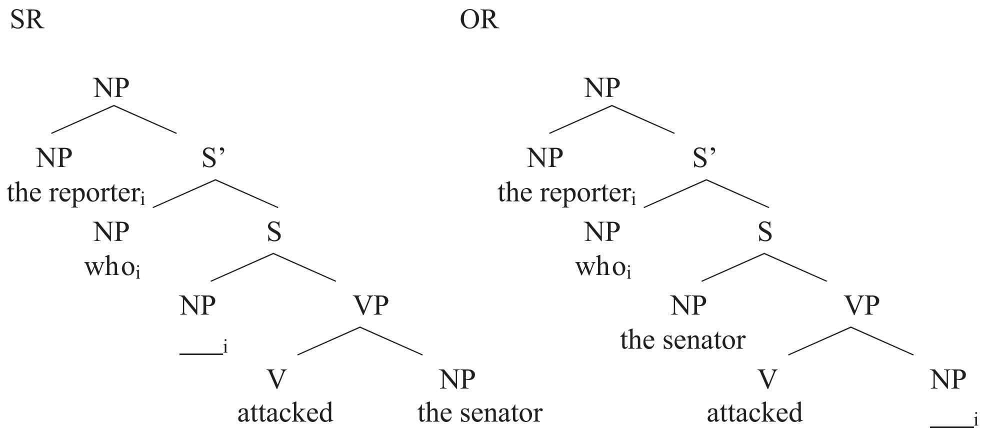
From a psycholinguistic perspective, King & Just (1991) and Just & Carpenter (1992) centered the debate on the working memory (WM) capacities of parsers: participants with low WM capacity had more difficulty in parsing and understanding object RCs. The question of how WM modulates language processing, along with the characteristics of object RCs that make them more costly for parsers, have been the subject of a large debate that has attracted widespread attention.

Several accounts have arisen, with two major groups of theories: (1) accounts that explain the subject RC-object RC asymmetry in terms of filler-gap dependency and; (2) accounts that focus on the frequency of use or the semantic features of the sentence (see Kahraman & Sakai, 2015 for a review of both perspectives).

The first group of accounts explains this asymmetry based on the greater difficulty to integrate information during processing in object RCs due to (a) greater structural complexity (or structural distance) between filler and gap (O'Grady, 1997; Ueno & Garnsey, 2008; Miyamoto & Nakamura, 2003) (as in example 2 below), or (b) longer linear distance between filler and gap (Linear distance accounts, like the Dependency Locality Theory (Gibson, 1998; 2000)) (see example 3):

(2)

Structural distance between subject RCs and object RCs (syntactic tree from Ueno & Garnsey, 2008):



(3)

Linear distance between subject RCs and object RCs (examples from Gibson, 1998):

a. Subject RC: [The senator$_i$] [that e$_i$ attacked the reporter] admitted the error.

Filler ⟷ Gap

b. Object RC: [The senator$_i$] [that the reporter attacked e$_i$] admitted the error.

Filler ⟷ Gap

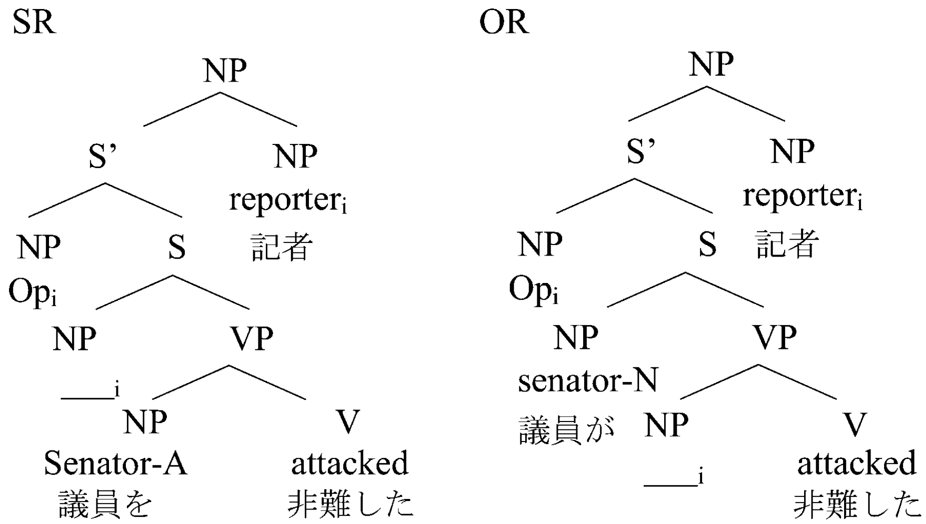
The linear distance hypothesis is based on the number of words that mediate between filler and gap. Longer dependencies, like the ones posited by object RCs, entail greater integration costs. Thus, parsers find it easier to integrate and process local dependencies: sentences in which the dependent element is located closer to its filler (Gibson, 1998; 2000).

On the other hand, the structural distance hypothesis posits that processing costs depend on the structural distance between filler and gap. In other words, in the number of syntactical nodes that mediates between both (O’Grady, 1997).

Both theories expect a greater processing cost, and greater WM burden for object RCs, in head-initial languages like English (as can be seen in (2) and (3)).

However, head final languages like Japanese predict different outcomes in both theories, as can be seen in (4) and (5).

(4) Structural distance between subject RCs and object RCs in Japanese (from Ueno & Garnsey, 2008):



(5) Linear distance between subject RCs and object RCs in Japanese:

a. Subject RC: [e$_i$ 議員を避難した] [記者 $_i$]

[e$_i$ Giin-o hinan shita] [kisha$_i$]

[e$_i$ Senator-ACC attacket] [reporter$_i$]

Gap ⟷ Filler

b. Object RC: [議員が e$_i$ 避難した] [記者 $_i$]

[Giin-ga e$_i$ hinan shita] [kisha$_i$]

[Senator-NOM e$_i$ attacket] [reporter$_i$]

Gap ⟷ Filler

As can be seen in 5, in Japanese the linear distance between filler and gap is shorter in object RCs than in subject RCs, reversing the pattern of head initial languages. However, object RCs exhibit greater structural distance between filler and gap, with more syntactic nodes between them than in subject RCs. For that reason, studies in this language were pivotal to distinguish between both accounts. Several authors (Ishizuka, 2005; Miyamoto & Nakamura, 2003; Ueno & Garnsey, 2008; Kahraman, 2012) have put these hypotheses to test, consistently showing that object RCs are more costly to Japanese speakers as well, thus supporting the structural distance hypothesis.

However, the heavier memory burden of object RCs is released when the embedded noun is a pronominal, as in sentence (6). This is explained as a result of the accessibility of first and second persons, which contributes to an easier integration (Warren & Gibson, 2002). The similarity between two full NPs is also stated as a factor that could hinder integration, making object RCs more costly to process by way of interference of the second NP (i.e. the subject of the embedded clause) (Lewis & Vasishth, 2005).

Effectively, it seems that this asymmetry is mediated by the accessibility and the competition between both nouns involved in the RC (Gennari & MacDonald, 2008). Moreover, the distributional frequency of the different structures has been also stated as a possible explanation for this asymmetry (Reali & Christiansen, 2007). These authors performed a corpus analysis and found that subject RCs are more frequent than object RCs, but this is only true when the noun phrases of subject and object are both explicit (like in the examples presented in (2) and (3)). However, when there is a personal pronoun in the embedded clause instead of a full noun phrase, like in (6), the frequency patterns are reversed.

(6) [The reporter$_i$] [that you attacked e$_i$] admitted the error.

Additionally, Reali & Christiansen found that these distributions had a reflection in the processing difficulties that parsers encounter: when the embedded noun was a personal pronoun, processing cost was higher for subject RCs than object RCs (see Christiansen & Chater, 2016, for a review of additional evidence for this view). However, Sato, Kahraman & Sakai (2010) showed that the mere frequency of subject RCs and object RCs cannot be the explanation for this asymmetry, at least in Japanese. These authors analyzed a written corpus of Japanese, and observed that subject RCs are not overall more frequent than object RCs (if any, object RCs were slightly more common).

Nevertheless, not only the use of personal pronouns seems to interact with asymmetry effects, but also it appears that the prototypicality of a given element (in particular, animacy) helps parsing and reduces processing costs. Thus, Traxler *et al.* (2002), Traxler, Williams, Blozis & Morris (2005) found that not only prototypical animate nouns help processing RCs (e.g. "A policeman arresting a thief" is more prototypical than the opposite, and thus this combination helps integration), but also importantly, they found that RCs with animate subjects and inanimate objects, like the ones in (7) and (8), reduce processing differences between subject and object RCs.

(7) Subject RC: The musician that witnessed the accident phoned the police.

(8) Object RC: The accident that the musician witnessed caused a number of injuries.

These results are not only due to similarity, since RCs which are also prototypical but where the subject was inanimate and the object animate did not show the same effect. Moreover, Mak, Vonk & Schriefers (2002; 2006) proposed that when both the HN and the embedded noun are bad prototypes of the subject, processing is not facilitated. Thus, it is precisely the combination of animate agents and inanimate patients which facilitates processing. These authors propose that parsers do not assign right away a role to a noun, but keep it in working memory, using animacy combinations to help integration and avoid reanalysis costs that could result from an early commitment to a subject RC interpretation. Similary, Gennari & MacDonald (2008) propose that object RCs with animate HNs are more costly because they entail a greater indeterminacy than object RCs with inanimate HNs. Upon appearance of inanimate HNs, parsers tend to assign them the recipient (or patient) role, while in the case of animate HNs, there is more variety in the possible interpretations entertained by parsers, showing a greater ambiguity. Semantic indeterminacy is closely related to the distributional patterns of the language as well. Sato, Kahraman & Sakai (2012) confirmed this tendency with Japanese too: in a corpus study, they observed that subject RCs generally occurred with animate HNs, while inanimate HNs were more likely to concur in object RCs (see Del Río, 2009; and Lin & Garnsey, 2011; Kahraman & Sakai, 2015, for a review of studies on the asymmetry of RC processing in head-initial and head-final languages, respectively).

Another proposal comes from MacWhinney & Pleh (1988), who focused on the number of changes of perspective involved in object RCs vs. subject RCs. In the case of subject RCs, there is no shift of perspective, since the subject of the matrix clause, which is also the element being talked about, is also the subject of the subordinate clause. However, in the case of object RCs there arises the need to shift perspectives and buffer the information prior to finishing the sentence. In object RCs that modify the subject of the matrix clause, the subject and topic of the main clause is not the subject of the subordinate clause; thus, parsers are forced to change the perspective from the subject of the matrix clause to the subject of the subordinate clause, and then back to the subject of the main clause after finishing the subordinate clause.

As it has been noted along the preceding pages, it is still not clear where the processing differences between object RCs and subject RCs lie. Both structural and semantic (and also pragmatic) factors interact to form a complex scenario that impinges on the processing costs parsers face. The study of head final languages would likely allow a better understanding of the (supposedly universal) mechanisms that may underlie the different processing costs. The distribution of subject and object RCs is tightly correlated with the production of these structures (Gennari & MacDonald, 2009; MacDonald, 2013a; 2013b). As will be reviewed in chapter 2, the same effects of animacy, namely, the tendency to assign animate elements to the agent position, has been reported in speech production studies. In this dissertation, we will explore the production of RCs, as it entails a structure that will help us explore the mechanisms underlying these universal differences, along with cross-linguistic variations.

In what follows, we will briefly review the main characteristics of Spanish and Japanese relative clauses from a descriptive standpoint, focusing our analysis on the two highest levels of the hierarchy, which will be topic of this PhD thesis: subject and object RCs.

Spanish Relative Clauses

In Spanish, just like in English, relative sentences are formed by an antecedent (an NP in the matrix clause) followed by a relativization marker and the subordinate clause. Main relativization markers in Spanish, classified as relative pronouns, are “que” (that/who/whom), “quien/quienes” (who/whom), “cuanto/ cuantos” (as much/many as), “cual/cuales” (which), “cuyo/a/os/as” (whose), “donde” (where), “cuando” (when) and “como” (how) (RAE, 2009: 1957). The most common and widely used marker (used with Subject, Direct Object, Indirect object, and Prepositional Object functions) is “que”, the marker that will be the focus of our study.

In general, descriptive accounts of Spanish grammar (e.g. Bosque & Demonte, 1999) consider that the relativization marker plays a syntactic function inside the sentence, acting as the subject (in Subject RCs) or the object (in Object RCs), as can be seen in examples (9a) and (9b), in which “que” is indexed with the antecedent “las personas” (“the persons”), acting as a pronoun (placeholder) for the NP “las personas” in the subordinate clause (the clause within brackets) and fulfilling the subject (9a) or direct object (9b) function (examples adapted from Fernández Lagunilla & Anula, 1995):

(9)

(a) Subject RC:

$[_{SN}$Las personas$_i$ $[_{OR}$que$_i$ conocieron a Juan]] venían de muy lejos.

$[_{NP}$The persons$_i$ $[_{RC}$that$_i$ knew Juan]] came from far away.

“The people who knew Juan came from far away”.

(b) Object RC:

$[_{SN}$Las personas$_i$ $[_{OR}$que$_i$ Juan conoció]] venían de muy lejos.

$[_{NP}$The persons$_i$ $[_{RC}$that$_i$ Juan knew]] came from far away.

“The people who Juan knew came from far away”.

However, some linguists (e.g. Demonte & Fernández Soriano, 2007; Sanz, 2005, for Spanish; Huddleston & Pullum, 2002, for English) have proposed that the relativization marker in these languages is not a “relative pronoun”, but a subordination marker, parallel to the one found in complement clauses (e.g. I said *that*, I think *that*…). In other words, relative clauses and complement clauses are marked by the same subordination item. *Que* (*that*) is a unique complementizing marker that signals the beginning of the subordinate clause, either with an adjectival (relative clause) or nominal (complement clause) function (Sanz, 2005; Sanz & Igoa, 2012), as can be observed in (10):

(10)

(a) Relative clause: El hombre **que** corría la maratón es mi primo.

The man *that* was running the marathon is my cousin.

(b) Complement clause: Creo **que** él corría la maratón.

I think *that* he was running the marathon.

Sanz and Igoa (2012) state that the reason why traditionally the complementizer has been thought to behave as a pronoun is because the subordinate verb is inflected, both in person and number. However, this fact only points to the existence of syntactic movement inside the sentence that creates a gap-filler dependence in relative clauses. Thus, the subordination marker is not the word co-indexed with the antecedent, but an empty category resulting from this syntactic movement, as can be seen if we analyze the example presented in (10a): (11):

(11) El hombre$_i$ [que e$_i$ corría la maratón] es mi primo.

The man *that* was running the marathon is my cousin.

Nevertheless, the distinction between different types of subordination[1] has a formal correlate in the sense that each of them presents its own morphological characteristics, given that "que" is not the only conjunction that can precede subordinate clauses (Fernández Lagunilla & Anula, 1995).

In addition to the above described formal characteristics of RCs, there are, broadly speaking two categories of relative clauses in Spanish in terms of their semantic function: *restrictive* or *specificative* and *appositive* or *explicative* relative clauses (RAE, 2009):

1. *Restrictive* or *Specificative*: The first group is composed of RCs that are used to identify, delimit and define the antecedent, by adding information that cannot be eliminated without changing the intended meaning. For example (from RAE, 2009: 3320),

(12) La casa *que* está en ruinas será destruida.

The house that is in ruins will be destroyed.

The sentence in (12) implies there are other houses, but only one, the one that is in ruins, will be destroyed, thus restricting the focus of the main clause.

[1] Not only nominal and adjectival, but also adverbial (Fernández Lagunilla & Anula, 1995).

2. *Appositive* or *Explicative*: In contrast to the first group, appositive RCs do not delimit the antecedent, but instead add extra information about it. These RCs are incidental modifiers, added by way of a digression. Consequently, they can be omitted without altering the meaning of the main clause. For example (from RAE, 2009: 3320),

(13) La casa, *que* está en ruinas, será destruida.

The house, which is in ruins, will be destroyed.

In this case, sentence (13) implies there is only one house, and that house will be destroyed, regardless of its condition of being in ruins. Thus, the clause between commas only provides additional information about the context in which the sentence appears.

Note that in Spanish the only difference between both RCs types is the presence of commas delimiting the subordinate clause, in written texts, and different prosodic patterns in oral speech: both types use the same complementizer *que*. However, in English, appositive RCs are marked by a different relativization marker: *that* (equivalent to *que*, in Spanish) in restrictive sentences, but *which* (or *who(m)*, depending on the antecedent) in appositive RCs.

We will focus on restrictive RCs in this study, due to their predominance in language and their function of adding new meaning. More importantly, these structures were selected for their similarities with Japanese RCs, as will be seen below.

Restrictive RCs in Spanish can be constructed until the lowest level of the NPAH, the object of comparison, as it was mentioned above. However, we will focus here solely on subject and object relative clauses (9a and 9b above). As can be seen in (14), in Spanish, object RCs allow the subject to be placed before the verb (a) or postponed to the end of the clause (b). The latter is actually the unmarked order for Spanish (Gutiérrez-Bravo, 2003).

(14) Object RCs:

a. Pre-nominal subject:

$[_{SN}$Las personas $[_{OR}$que Juan conoció]] venían de muy lejos.

$[_{NP}$The persons $[_{RC}$that Juan knew]] came from far away.

"The persons who Juan knew came from far away"

b. Post-nominal subject:

[$_{\text{SN}}$Las personas [$_{\text{OR}}$que conoció Juan]] venían de muy lejos.

[$_{\text{NP}}$The persons [$_{\text{RC}}$that knew Juan]] came from far away.

"The persons who(m) Juan knew came from far away"

In spite of the fact that Spanish allows flexible word order, RCs do not allow more variation than these two positions in object RCs. When the agent is the HN, the information can only be expressed with the same word order as was seen in (9a): HN [verb – object], in the active form, resulting in a subject RC. On the other hand, when the patient is the HN, voice variation is allowed, but the resulting word order will remain the same, as can be seen in (15):

(15)

a. Patient as HN – object RC with active verb:

La chica [$_{\text{CR}}$que mira el anciano] lleva un vestido rojo.

The girl [$_{\text{RC}}$that looks the old man] wears a red dress.

"The girl who the old man is looking at is wearing a red dress".

b. Patient as HN – subject RC with passive verb:

La chica [$_{\text{CR}}$que está siendo mirada por el anciano] lleva un vestido rojo.

The girl [$_{\text{RC}}$that is being looked at by the old man] wears a red dress.

"The girl who is being looked at by the old man is wearing a red dress".

Japanese Relative Clauses

As a head-final language[2], in Japanese relative clauses are presented in prenominal position. Thus, relative clauses in Japanese precede their antecedents (which should be properly called 'postcedents') (from Tsujimura, 2007):

(16)

a. Satoo-sensei-ga [gakusei-ga kaita] ronbun-o yondeiru.

[2] In head-final languages, verbs are located at the end of the sentence, and complements and modifiers are placed prior to the phrasal (modified) head.

Prof. Sato-NOM [student-NOM wrote] article-ACC is reading.

Prof. Sato is reading the article that the student wrote.

b. Hanako-ga [Taroo-ga kaita] hon-o yonda.

Hanako-NOM [Taro-NOM wrote] book-ACC read.

Hanako read the book that Taro wrote.

On the other hand, as can be observed in the previous examples, and in contrast with Spanish, there are no relative pronouns (relativization markers) in Japanese. The subordinate verb and the antecedent of the clause are joined directly, with neither morphological nor syntactic overt signal that indicates whether the sentence is an RC or not until the HN, in the main clause, is uttered (Tsujimura, 2007), thus creating a temporary ambiguity (Hirose, 2006).

Similarly to Spanish, the most basic types or relative clauses are subject RCs and object RCs (Iori, 2001) (along with other types down the NPAH hierarchy (Iwasaki, 2002)), and once again these two will be the focus of our research. The only difference between subject RCs and object RCs rests on the case particle that accompanies the HN in the subordinate clause[3], with no changes in the linear ordering of constituents:

(17) Subject RC:

[e_i Doroboo-o utta] keisatsukan$_i$-wa rainen teinen da.

[Thief-ACC attack-past] policeman-TOP next year retirement is.

“The policeman who attacked the thief will retire next year”.

(18) Object RC:

[Doroboo-ga e_i utta] keisatsukan$_i$-wa rainen teinen da.

[Thief-NOM attack-past] policeman-TOP next year retirement is.

“The policeman who the thief attacked will retire next year”.

---

[3] Japanese has a rich system of morphological case marking, which allows listeners to determine agent, receptor, dative, etc. before the verb of the sentence. In particular, in relative clauses, the particle *o* (accusative marker) is used in subject relative clauses, since the subordinate NP acts as the object of the clause. Similarly, the particle *ga* (nominative marker) will be used in object RCs, as the subordinate NP is the subject of the embedded clause.

Similarly, when sentences with the patient as HN, like the one in (18), are expressed with a passive voice, there is no change in word order, but only in case particle and verb conjugation:

(19) Patient as HN – subject RC with passive voice:

[Doroboo-ni e_i utareta] keisatsukan$_i$-wa rainen teinen da.

[Thief-DAT attack-passive] policeman-TOP next year retirement is.

"The policeman who was attacked by the thief will retire next year".

Nevertheless, there are several differences with Spanish regarding the categories of RCs:

In the first place, the syntactic structure of relative clauses (17) to (19) is the result of a syntactic movement, which gives rise to a gap of the antecedent inside the subordinate clause, marked with the letter "e" and indexed with its corresponding filler (*keisatsukan* – policeman). Thus, as can be seen in the examples above, in Japanese the *filler* of the relative clause is located before its *gap*. However, this order is reversed in Spanish (and English):

(20) El policía$_i$ [que e_i atacó al ladrón] se jubilará el año que viene

The policeman [that attack-past to the thief] retire-fut next year.

"The policeman who attacked the thief will retire next year".

However, there is another type of relative clauses in Japanese that does not correspond to any type of RC in Spanish (or English): RCs with no *gaps,* also called "attributive sentences" (Comrie, 2002; Ozeki & Shirai, 2007). In this type of RCs, there is no syntactic movement (no gap-filler dependency), as shown in examples 21 to 23 (from Tsujimura, 2007):

(21) [Musuko-ga iede-shita] Taroo

[Son-NOM ran away from home] Taro

Taro whose son ran away from home

(22) [Meijin-ga ryoori-shita] aji.

[Expert-NOM cooked] flavor.

Flavor (that results when) an expert cooked.

(23) [Hito-ga tooru] monooto.

[Person-NOM pass] sound.

"The sound of people going by".

The nouns that follow the relative clauses (inside the brackets in these examples) are their head nouns. In this type of sentences the subordinate clause stands as a complete sentence by itself. The dependence with the HN in this type of sentences is not determined in a syntactic manner, but rather the relation of subordination is established through semantic and pragmatic bonds (Tsujimura, 2007). This type of sentences do not present a gap-filler dependency. This fact has led to the idea that RCs in Japanese (and other East Asian languages, like Korean) are qualitatively different from RCs in European languages, like Spanish, due to their semantic characterization, in contrast to the syntactic characterization of Spanish RCs (Ozeki & Shirai, 2007)[4].

Secondly, the main distinction in Spanish RCs described above between appositive and restrictive RCs is not expressed in Japanese (Yamada, 1995). Thus, the sentence in (24a) with a restrictive RC has its corresponding counterpart by means of an RC in Japanese as well, as can be seen in (24b) (Examples from Yamada, 1995; English translation ours):

(24)

(a) Los niños que viven lejos llegan tarde al colegio.

The children who live far arrive late at school

(b) Tooku-ni sundeiru kodomo-tachi-wa gakkoo-ni chikoku-suru.

Far-in live-ger-be-pres child-pl-TOP school-in delay-do-pres.

"Children who live far are late to school".

[4] Specifically, Ozeki & Shirai note that Japanese RCs, especially the types presented in (19) to (21) have a series of particular features that make them follow a different developmental pattern in the interlanguage of L2 learners from that observed in relative clauses with gap-filler relation. This pattern does not follow the NPAH (Keenan & Comrie, 1977), leading authors to hypothesize that the underlying structures might differ.

However, this is not the case for the appositive sentences. The Spanish appositive RC in (25a) cannot be expressed in the same manner in Japanese, where it must be conveyed by means of a coordinate sentence instead (25b)[5].

(25)

(a) Los niños, que viven lejos, llegan tarde a la escuela.

The child-pl, that live far, arrive late to school.

Children, who live far, are late to school.

(b) Sono kodomo-tachi-wa tooku-ni sundeite, sorede gakkoo-ni chikoku-suru.

Those child-pl-TOP far-in live-progr-te, therefore school-in delay-do-pres.

Those children live far, and therefore are late to school.

For this reason, as stated at the end of previous section, we will concentrate on restrictive RCs in the current study.

General summary

In this study we will analyze the interplay of relational and non-relational processes in monolingual and bilingual sentence production through the use of RCs in two typologically different languages: Spanish and Japanese, with different animacy combinations. Although RCs are universally shared (Keenan & Comrie, 1979), there are some differences in the way RCs are constructed in these two languages, as was described along the previous pages. The main points of this review to take into account are as follows:

---

[5] However, Iori (2001) shows some cases in which appositive clauses can be used in Japanese, with the use of proper nouns or definite NPs with demonstratives like "this", "that", etc., like the ones presented in (1) and (2) (from Iori, 2001):

(1) Kino atta Tanaka-san.
Yesterday met Mr. Tanaka.
Mr. Tanaka, whom I met yesterday.

(2) Taro-ga yomitagatteiru kono hon.
Taro-NOM want to read this book.
This book, which Taro wants to read.

This type of RC is not commonly used when referred to common nouns, so we decided to exclude it from the present study.

1. Subject and object RCs are both allowed in Japanese and Spanish. The functions they fulfill are the same in both cases.

2. Although there is no complete consensus, the syntactic structure of subject and object RCs in both Spanish and Japanese is also shared: RCs in both languages are the result of syntactic movements that entail gap-filler dependencies[6].

3. The order of constituents in both languages is reversed. In Spanish the HN (*filler*) appears first, acting as a guide to process subsequent information. In Japanese, the HN does not appear until the relative clause has been processed; the disambiguation of the input is done mainly by means of case particles.

4. The linear order of the elements in relative clauses in Japanese does not vary. The only difference between subject and object RCs is the particle used in each case. When the agent is the HN, the active voice is the only accepted form. However, when the patient is the HN, RCs can be expressed in active or passive voice, with no changes in word order, but only in case particles and verbal form.

5. Similarly, in Spanish the order of the elements allows little variation: in subject RCs, the subordinate NP (the object of the verb) is always placed after the verb; however, object RCs allows both pre and post-verbal subjects (i.e. the subordinate NP in these structures). Same as in Japanese, RCs in Spanish with the agent as HN only allow subject RCs in active form, while RCs with the patient as HN allow for both active and passive forms (the former resulting in an object RC, and the latter in a subject RC), being the passive form identical in word order to object RCs with post-verbal subject.

[6] Although, as was reviewed earlier in this chapter, there is a type of RCs that lacks gap-filler dependencies (i.e. so-called attributive sentences), which have different functions from Subject or Object RCs.

Chapter 2. Monolingual sentence production

In order to articulate a message, a speaker must undergo a series of stages from the conceptual representation of the message to its final articulation (Bock & Levelt, 1994; Ferreira, 2010). Although there are differences among authors, it is generally agreed that the processor is composed of three separate parts each playing a different role (Bock & Levelt, 1994; Levelt, 1989; Ferreira, 2010; Ferreira & Slevc, 2007):

1. *Message encoding* processes (Ferreira & Slevc, 2007; Ferreira, 2010), or *conceptualizer* (Levelt, 1989), encode a non-linguistic representation of the concept to be communicated, a yet pre-linguistic rendering of the message to be uttered.

2. *Grammatical encoding*: Once the pre-linguistic message has been created, the central phases of the linguistic formulation process take place. It is here that the translation of the conceptual message into a linguistic representation to be finally articulated takes place. This component is broken down, in turn, into various sub-processes that will be explained below.

3. *Articulation processes* (Ferreira & Slevc, 2007) or the *articulator* (Levelt, 1989): This component is intended to perform a set of motor instructions from the phonological form of the message for its final articulation.

Grammatical encoding processes are in charge of the elaboration of the linguistic message that will be uttered. As mentioned, it is the central part of the process and the component that has received a greater amount of attention in psycholinguistic research. This phase is divided into two sub-components, which proceed in a more or less parallel fashion: *Lexical Encoding* (Garrett, 1975, 1982) or *Content Processing / Processes* (Ferreira & Slevc, 2007; Ferreira, 2010) and *Grammatical Encoding* (Garrett, 1975, 1982) or *Structure Building / Processes* (Ferreira & Slevc, 2007; Ferreira, 2010). Each of these processes takes place, in turn, in a subsequent series of steps, as can be observed in Figure 2.1.

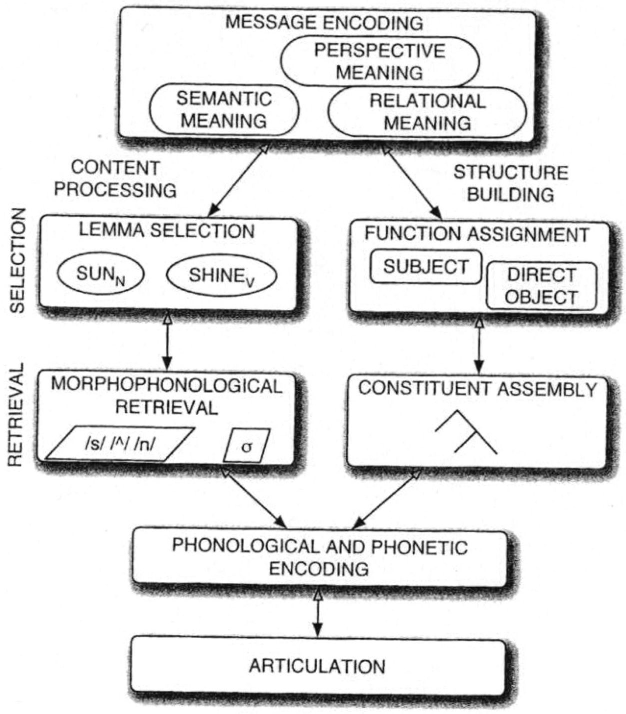


Figure 2.1. General structure of the language production system (Ferreira & Slevc, 2007).

First, during *Content processing* (on the left in the figure), the lexical items to be included in the uttered message are selected. These elements are called *lemmas*. *Lemmas* are pieces of information that contain both the meaning and all the grammatical features of each individual lexical concept, as well as its relevant use. For example, verb lemmas contain information concerning the arguments they take and their relations. Lemmas lack, nonetheless, phonological form, which is encoded in the following step.

At the same time, *Structure Building* occurs, once again at two different levels: the functional level (*Function Assignment*) and the positional level (*Constituent Assembly*). In the former, elements retrieved from the mental lexicon (*lemmas*) are assigned the grammatical roles they will fulfill in the final utterance: subject, object, etc., and a structural scaffold of the message that will be uttered is created (Bock & Ferreira, 2014). At the positional level, fixation of the order of the elements is undertaken. This linearization process results in the final form of the utterance, and is closely related to the previous functional level.

Research on sentence production mechanisms has focused on Structure Building processes (the right side of Figure 2.1) and in the mapping process that takes place between lexical (semantic) selection and grammatical encoding processes. In general, there is a close relationship between functional relations and grammatical function assignment. In this way, agentive elements (or agents of the event) are usually assigned the subject (nominative) role and position, themes are given the object (accusative) role and position, oblique arguments (e.g. recipients) the indirect object role and position, and so on, although in some cases, like passive sentences, this relation can be reversed. The selection of one structure over the other may be due to several discursive factors, or merely to differential accessibility reasons.

Even though production processes take place in a series of steps, the whole process is developed in an extremely fast and automatic way; in other words, linguistic production is characterized by its highly incremental character (Levelt, 1989; Levelt, Roelofs & Meyer, 1999). In this sense, the production system does not have to wait until all the information is made available at one given stage to start with the following phase, thus allowing the fluent speech that characterizes native speakers of a language. This fact allows the system to be sensitive to the relative accessibility of the elements involved in the utterance: most accessible elements are retrieved earlier and, therefore, are processed and produced earlier than less accessible elements, thus enabling fluent production. Each of the levels inside grammatical encoding can be influenced by different types of information. Thus, the functional assignment level can be influenced by conceptual information, such as animacy (e.g. Branigan, Pickering & Tanaka, 2008; Gennari, Mirkovic & MacDonald, 2012) or imageability (Bock & Warren, 1985); the positional level, in turn, can be influenced by lexical and syntactic information (Tanaka, Branigan, McLean, & Pickering, 2011), and, as some studies suggest, under certain circumstances, it can receive influence from conceptual information as well (e.g. Branigan & Feleki, 1999; Prat-Sala & Branigan, 2000).

Moreover, many studies have embarked on the task of exploring the timing in which accessible elements or structures are selected, thus affecting the final utterance. They explore the moment during speech planning at which speakers rely more on lexical information or on structural information, how both types of information are used during the whole planning process, and which of them predominates when deciding the final utterance. Eye-tracking studies and planning scope studies have focused on these issues.

In what follows, we will first review the effects of accessibility, both conceptual and structural, on sentence production. After that, we will review eye-tracking and planning scope studies. We will introduce cross-linguistic differences in incremental processing that have turned out to be essential for an understanding of the complexities of speech planning.

Conceptual accessibility

Accessibility has repercussions in the information used in both function assignment and constituent assembly processes, with most accessible elements bearing more dominant grammatical roles, or appearing earlier in the utterance than less accessible ones. The way conceptual information intertwines with both sub-processes has been widely studied and debated, as a way to assess how both of them develop.

The first subcomponent, function assignment, is in charge of assigning a grammatical function to each of the elements of the message. This assignment is established incrementally and in accordance with the Noun Phrase Accessibility Hierarchy (Keenan & Comrie, 1977), such that the first assigned element will be the subject, followed by the direct object, the indirect object, etc. (see chapter 1 for an explanation of the NPAH). Subsequently, in this process conceptually more accessible elements are assigned their grammatical function first, and, consequently, will receive the subject role more frequently.

During constituent assembly processes, the elements that will form part of the utterance will receive their linear order. Similarly to the preceding stage, more readily accessible elements will be allocated first, occupying the initial position of the to-be-uttered sentence (Bock & Warren, 1985; Kelly, Bock & Keil, 1986; McDonald, Bock & Kelly, 1993; Onishi, Murphy & Bock, 2008; Tanaka *et al.*, 2011).

Traditional Serial Models (e.g. Bock & Warren, 1985; McDonald, *et al.*, 1993), have assumed a completely sequential process (with no overlap between stages). These studies have manipulated the conceptual accessibility of the elements in an event, and, subsequently, have analyzed the linguistic production of different kinds of structures. Among conceptual accessibility factors, animacy has been one of the most widely studied, due to its character as an inherent feature of the elements, and its fairly universal nature across languages. However, this is not the only factor that has been explored: others are discourse focus, imageability, concreteness or prototypicality. (e.g. concreteness / imageability: Bock & Warren, 1985; prototypicality: Kelly, *et al.*, 1986).

Bock and Warren (1985), for example, analyzed the role of imageability in production, by means of a study of sentence recall with (active / passive) declarative and (prepositional / double-object) dative sentences, in comparison with phrasal conjunct sentences: the former two constructions involved changes in functional role along with change in word order, whereas the latter involves only word order change. Examples of these sentences are given in (1), (2) and (3) (from Bock & Warren, 1985):

(1) Simple declarative

(a) Active: The doctor administered the shock.

(b) Passive: The shock was administered by the doctor.

(2) Dative

(a) Prepositional: The old hermit left the property to the university.

(b) Double object: The old hermit left the university the property.

(3) Phrasal conjunct

(a) Natural order: The lost hiker fought time and winter.

(b) Unnatural order: The lost hiker fought winter and time.

Results showed a tendency to reverse word order with the more imageable element being placed earlier, both in declarative and dative sentences[7]. However, phrasal conjunct elements did not show the same tendency, with the same number of inversions (very few in general) regardless of imageability. Hence, these

[7] In other words, active declarative sentences with less imageable elements first were recalled as passive sentences and vice versa. As for dative sentences, prepositional sentences with the less imageable element placed earlier were remembered as double-object sentences, and vice versa.

authors concluded that conceptual accessibility affects functional assignment, and not constituent assembly, as word order was the only difference between the first two types of structures and the last one.

Similarly, Bock, Loebell and Morey (1992) and McDonald, *et al.* (1993) provided more evidence in favor of this proposal, by using animacy as a measure of conceptual accessibility. Bock *et al.* (1992) studied the mapping processes between conceptual information and functional assignment by means of a structural priming experiment. They used declarative sentences with either animate subject / inanimate object or inanimate subject / animate object as primes and recorded the type of structures that participants used to describe subsequently presented pictures. The results show, first of all, that there was a general tendency to use passive sentences when the agent of the action was inanimate, rather than the opposite pattern. This happened independently of priming effects. However, when primes were active sentences with inanimate subjects, the number of active inanimate-animate sentences increased. There was a priming effect for both structure (active sentences were produced more often after active primes and vice versa) and animacy (inanimate subjects were produced more often when the primes had inanimate subjects and vice versa). In conclusion, these results point to a primacy of animacy in the assignment of grammatical functions in the sentence, which could also be influenced by different accessibility effects (i.e. mediated by priming effects).

McDonald *et al.* (1993) also examined the role of animacy (along with other variables such as word length and prosody) in declarative sentences. In this case, they compared the performance with declaratives with that in sentences with phrasal conjunctions, by means of a sentence recall task. Their results replicated the pattern found by Bock and Warren (1985) for imageability and the previously described results of Bock *et al.* (1992) with animacy, as there was a greater number of inversions consisting of raising the animate element to subject position, but only in declarative sentences and not in phrasal conjunctions[8].

[8] This study also showed that accessibility is sensitive to semantic features, as opposed to lexical properties: there was not an influence of word length on the number of inversions. Thus, the locus of accessibility is conceptual, not lexical (McDonald, *et al.*, 1993).

However, these early studies focused only on English, a language in which the grammatical subject is placed almost always at the beginning of the sentence, due to the strict word order that characterizes it. This makes it difficult to disentangle the effects of both subprocesses in the final shape of the utterance (Branigan & Feleki, 1999). Therefore, in recent years, there has been a rapid increase of investigations that examine the role of animacy (or other conceptual factors) in production in languages whose canonical order differs from that of English in that they have less stringent serial order rules (e.g. Greek: Branigan & Feleki, 1999; Spanish: Prat-Sala & Branigan, 2000; Japanese: Tanaka, *et al*., 2011; Branigan, Pickering & Tanaka, 2008; Serbian: Gennari, *et al.*, 2012). Results in these languages show that there is an influence of conceptual accessibility on the positional level, in some cases independently or in addition to grammatical function assignment. Both levels can overlap in time, rather than being completely sequential phases. Due to this overlapping activation, conceptual accessibility can even solely affect the positional level, such that more accessible elements can be placed earlier in the sentence, regardless of their grammatical function (Prat-Sala & Branigan, 2000).

Branigan & Feleki (1999) compared direct transitive sentences in Greek with different animacy combinations[9], in a study with a sentence recall task. Their results showed that participants tend to collocate the animate element at initial position, regardless of its grammatical function: participants changed word order to OVS, to the detriment of the preferred SVO, when the object of the sentence was animate. The opposite tendency (i.e. changing the order to SVO when the subject was animate) was also present. Shortly after, Prat-Sala and Branigan (Prat-Sala & Branigan, 2000) continued this line of research begun with Greek, looking closely at the influences of conceptual accessibility on various languages. In particular, these authors also carried out a study in Spanish, a language with relatively flexible word order[10], and compared their results with the English data. These authors manipulated contextual saliency (an extrinsic element that favors accessibility: "derived accessibility" in the authors' words), in interaction with

[9] Greek is a language that holds a case marking system, and that, consequently, accepts word orders like OVS, although the canonical order is SVO (Branigan & Feleki, 1999).

[10] Unlike Greek, Spanish does not make use of case marking. Flexibility of sentence elements is possible thanks to a rich verbal morphology system.

animacy (an intrinsic element that favors accessibility: "inherent accessibility"). Participants in this study, native speakers of either Spanish or English, had to describe a scene, in which an agent performed a transitive action over a patient. In half of the trials, previous context had made the agent more salient and, in the other half, the patient was made salient (derived accessibility). In turn, in the first study both agent and patient were always inanimate, whilst in the second, the agent was always inanimate and the patient animate. The results showed that in both experiments, and regardless of language, speakers have a tendency to assign more prominent syntactic positions to the elements made more accessible by the previous context, either agents or patients of the action. However, in the case of Spanish, in contrast to English, the prominence was achieved not only through the use of passives, but through the use of dislocated active sentences, of the type in (4), to the same extent as passives.

(4) A la mujer$_i$ la$_i$ atropelló el tren.

to the woman$_i$ her$_i$ ran over the train.

The woman, the train ran over ("The train ran over the woman").

English speakers, in general, showed a stronger tendency to produce non-typical structures than Spanish speakers, thus preferring to stick with the salient element as the subject of the sentence, independently of its bearing the agent or the patient role. However, Spanish speakers showed a tendency to both place the salient element at the beginning of the sentence, and changing its grammatical role, such that the salient element received the subject function. In experiment two, both inherent and derived accessibility summed their effects, so that an animate patient made salient led to more non-typical structures than an inanimate one, and to the same extent in both languages. Taken together, those studies seem to suggest that not only grammatical function can be an important feature in assigning a prominent syntactic position to salient elements, but word order alone can also have a significant role.

As for verb-final languages, evidence from Japanese comes from the work of Montag and MacDonald (2009). In this study, the authors used a picture-description paradigm to compare the production of object relative clauses by native speakers of Japanese and English, in which the object (i.e. the head noun) could be either animate or inanimate. Their results showed that there is a strong tendency to avoid object relative clauses with animate objects, both in English and Japanese. However, this tendency was much stronger in Japanese, showing a close relation between grammatical function and animacy in this language, despite its opposite word order. In this study, subjects were always animate (the first element presented in Japanese was always animate) and only the element at the end of the sentence (the head noun) had animacy variation. For example (from Montag & MacDonald, 2009):

(5) Active

(a) Animate-animate:

Onnanohito-ga nage-te-iru otokonohito.

Woman-NOM throw-Pres-Prog man

“The man (that) the woman is throwing”

(b) Animate-inanimate:

Onnanohito-ga nage-te-iru booru.

Woman-NOM throw-Pres-Prog ball

“The ball (that) the woman is throwing”

In Japanese, also, the construction of passive sentences inside relative clauses does not involve changes in word order, but only changes in grammatical function:

(6) Passive

Onnanohito-ni nage-rare-te-iru otokonohito

Woman-BY throw-Pass-Pres-Prog man

“The man being thrown by the woman”

Thus, this study supported the idea that word order does not play an important role in the mapping processes between syntax and semantics, as it is not affected by animacy: grammatical functions with non-prototypical animacy combinations (as a theme being an inanimate element) were avoided, using a different verbal voice instead.

However, evidence drawn from the work of Tanaka *et al.* (2011) seems to point to different conclusions. By using a sentence recall task, these authors tried to explore the role of animacy in sentence production in Japanese. They presented their participants with a series of direct transitive sentences in which the first element could be either an animate element or an inanimate element. At the same time, they presented two types of word orders (both with active verbs): SOV (canonical word order) and OSV (scrambled word order). Results showed that there was a tendency to change word order when the first noun was the inanimate object, to a greater extent that in the rest of combinations. When an animate object appeared at the beginning of the to-be-recalled sentence, there was also a considerable number of cases in which the order was changed, with no change in verb form, as well as some cases of changing the verbal voice. In general, there was a significant trend to keep canonical word order when recalling sentences, independently of animacy[11]. In a subsequent experiment, authors corroborated the tendency observed in their first experiment, that is, that grammatical function and word order are jointly collaborating: both levels played a role in the assignment of a more prominent role to accessible elements (i.e. more prominent role could be defined as either an element placed earlier in the sentence or an element assigned the subject function, or a combination of both). In this second experiment, the authors added a third condition: verbal voice (active or passive). Importantly, participants in this study tended to recall sentences with a different verbal voice when this resulted in the animate element taking the subject function, but this was done to a greater extent when the first shown element was the animate one (i.e. participants were not changing word order in these cases)[12]. In conclusion, both experiments by Tanaka and colleagues seem to suggest an influence of conceptual information at both levels, functional and positional, in a cooperative way.

---

[11] This was also found by Branigan and Feleki (1999) with modern Greek.

[12] For example, participants tended to recall the sentence (1a) as (1b) more frequently than when the inanimate element "booto-ga/o" appeared at the beginning of the sentence:

(1)

(a) Minato-de ryoshi-o booto-ga hakonda
Port-in fisherman-ACC boat-NOM carried.
"In the port, the boat carried the fisherman".

(b)Minato-de ryoshi-ga booto-ni hakobareta
Port-in fisherman-NOM boat-by was carried.
"In the port, the fisherman was carried by the boat".

Despite the fact that the participants in Montag and MacDonald's work were not affected by word order, we may say that in general, the conclusions of Montag & Macdonald's and Tanaka *et al*'s studies are not incompatible. The types of structures differ in both studies: Montag and MacDonald made use of a structure which allows limited word order movement, while at the same time they presented the animate element always at the beginning of the sentence. Tanaka *et al.*'s interaction of levels may be due to the possibility of free word order, together with different animacy combinations. Further studies are required in Japanese. Nonetheless, all these data, taken together, point to the fact that the way in which both levels add their effects (in case they do), as well as the weight of conceptual information on each level, is highly dependent on the language and its particular grammatical structures.

Additional evidence for this position (i.e. influence of conceptual accessibility on both levels, depending on language) comes from work by Gennari, *et al.* (2012). These authors compared the influence of conceptual accessibility in production processes across three typologically different languages: Serbian, Spanish and English. To do so, they used a picture description task with object relative clauses, in which the agent of the action (which is at the embedded position) was always animate, whereas the theme (head noun of the sentence) could be either animate or inanimate. Freedom in word order is greater in Spanish and Serbian[13] than in English, and there are impersonal structures in the first two languages that allow the omission of the subject in a sentence. These structures only mention the object, suggesting the existence of a general unknown (or unmentioned) subject. This kind of structure is not available in English in the active voice. For example (from Gennari *et al.*, 2012):

(7) Spanish Active impersonal:

El hombre (al) que están golpeando

the man who(m) are hitting

'The man being punched'

(8) Serbian Active impersonal

[13] Serbian is a language that uses a case marking system, thus different from Spanish. Both have relatively free word order, but their word order preferences differ.

Cˇ ovek koga udaraju

man-NOM whom are hitting

‘The man being punched’

Results showed that, due to this greater syntactic variability (and the usual tendency to avoid passive sentences by native speakers of Serbian and, to a lesser extent, of Spanish), both Serbian and Spanish speakers tended to produce impersonal sentences. This was the case when the theme of the action was animate. Participants constructed, instead, impersonal structures like (7) and (8), in which animate elements are kept as the object of the sentence, but the subject is omitted.

Therefore, the results from this study (in combination with Montag and MacDonald’s results with Japanese, previously reviewed), indicate that not only the animacy of the subject (the most prominent position) is important in the selection of the final structure: even when the agent remains animate, animate themes, non prototypical, also affect sentence structure choices.

In conclusion, conceptual accessibility can produce an effect in both grammatical function assignment (subject, object, etc.) and serial ordering. The weight of this influence on each of these levels, and whether they overlap during their activation, remains unknown. However, there seems to be evidence that the relative weight conceptual accessibility has in each of them is dependent on the grammatical structures provided by each language, as speakers make use of all the cues that their language supplies in order to fulfill the demands of a highly incremental production system.

Structural accessibility

As we have seen so far, conceptual factors play an important role in structure selection during the production process. However, these factors are not the only ones affecting the choice of syntactic structure. The accessibility of a previously selected (or listened) syntactic form also plays an essential role therein (e.g. Bock, 1986; Pickering & Branigan, 1998; Hartsuiker, Bernolet, Schoonbaert, Speybroeck & Vanderelst (2008); Hartsuiker & Westenberg, 2000; see Ferreira & Bock, 2006; Pickering & Ferreira, 2008, for a review; and Mahowald, James, Futrell & Gibson, 2016, for a recent meta-analysis). The previously reported conceptual accessibility has been shown to exert an influence on grammatical function assignment and/or positional order, thereby revealing a relation between semantic and syntactic information. Thus, the selection of syntactic structure is determined, not only by the elements that are more salient during message encoding, but also by canonicity factors, contextual saliency or discourse factors (Levelt, 1989). But in addition, the probability of appearance of a given syntactic form in a language also determines its degree of accessibility and, consequently, the degree of use in future utterances (Gennari & MacDonald, 2009). This means that previously selected syntactic structures will be more salient and will result, subsequently, in a greater degree of use in the following utterances. The existence of this "structural priming" effect, that is, the effect of structural accessibility on the choice of syntactic structure independently of words, provides the opportunity of telling apart structural planning from conceptualization processes (Dell & Ferreira, 2016).

Starting with Bock (1986), several studies have shown a tendency of speakers to repeat the same structure of previous utterances, independently of semantic factors (e.g. increased production of passives over actives –e.g., Bock, 1986; Bock, *et al.*, 1992; Ferreira, Bock, Wilson & Cohen, 2008; Hartsuiker *et al.*, 2008; Messenger, Branigan, McLean & Sorace, 2012–, or increased production of prepositional object over double object sentences –e.g. Bock, 1986; Ivanova, Pickering, Branigan, McLean & Costa, 2012; Cai, Pickering & Branigan, 2012; Bernolet, Collina & Hartsuiker, 2016). Importantly, evidence points to the fact that structural priming is not due to superficial similarities, but to the repetition of syntactic processes and/or representations, inasmuch as superficially similar but syntactically different constructions do not lead to structural priming effects: for instance, the sentence "Susan brought a book to study" does not prime a superficially similar sentence like "Susan brought a book to Stella" (Bock & Loebell, 1990). Even languages where syntactic information plays a secondary role in parsing (and planning), such as Mandarin, have shown an independence of structure selection from meaning (Huang, Pickering, Yang, Wang & Branigan, 2016).

The effects of structural priming have been explained in terms of residual activation of the previously used/listened constructions (Hartsuiker & Pickering, 2008): previously used linguistic elements are more active in the speaker's mind, as an effect of remaining activation, which is still present for a short period of time (Branigan, Pickering & Cleland, 1999). Thus, when the speaker comes across a message that can be mapped onto a linguistic structure that resembles the prime in its structural functions, the still activated construction will tend to be selected over others that are less activated. For example, when speakers have produced previously sentences like the one shown above (reproduced in (9)):

(9) Susan brought a book to Stella.

and they are faced with the task of constructing a sentence with the verb "give", for example, they will use (10) more often than (11), in comparison with a condition in which they were not primed.

(10) John gave a present to his mother.

(11) John gave his mother a present.

However, priming is not only present at short intervals, for also long lasting effects have been found. These effects have been found not only within the same experimental session with different amounts of intervening items in-between (e.g. Hartsuiker & Kolk, 1998; Branigan, Pickering, Stewart & McLean, 2000; Bock & Griffin, 2000; Bock, Dell, Chang & Onishi, 2007; Bernolet, Collina & Hartsuiker, 2016), but also after even one week, in the form of cumulative priming (e.g. Kaschak, 2007; Kaschak, Kutta & Schatschneider, 2011; Kaschak, Kutta & Coyle, 2014). Activation accounts have difficulties explaining these long lasting effects. Implicit learning accounts based on connections models (Chang, Dell & Bock, 2006; Chang, Dell, Bock & Griffin, 2000; Chang, 2002) offer and alternative explanation to these effects in terms of adaptations of the system due to mismatch between expectations and comprehended speech. This results in long lasting learning.

Structural priming has been observed in a different set of languages, being consistent not only in head-initial languages, but also in head-final languages, like Japanese. Tanaka, Pickering and Branigan (2009) analyzed the effect of syntactic accessibility in Japanese with a picture-matching/-description task preceded by a priming phase. They compared the production of Japanese speakers in canonical/ scrambled sentences (with no change in grammatical function but a different word order: SOV vs. OSV order) with the performance in active/passive sentences (which in Japanese do not involve a change in word order, but merely a change in grammatical function); both conditions were crossed with different animacy combinations (although the target picture was always agent-inanimate). They found structural priming in both word orders (with SOV sentences being more likely to be recalled as SOV, and OSV sentences being more likely to be recalled as OSV) and in verbal voice, with a tendency to repeat the same voice in uttered sentences.

As these results show, structural priming does not take place only at the level of grammatical function assignment, but also affects word order. Hartsuiker and Westenberg (2000) provide with evidence toward this conclusion. In their study, with Dutch native speakers, they made use of a sentence completion task. These authors tried, as their main purpose, to show that the influence of syntactic accessibility is located at (or at least, can affect) the positional level (or "linearization"), as a way to allow the speaker to deal with word order variations, thus fulfilling fluency requirements. They used subordinate clauses that required the use of an auxiliary and a participle for completion[14]. Their results show that there was a priming effect even in this kind of structures is which grammatical function does not vary, thus providing strong evidence for the effect of structural accessibility on word order. Our own research team (Rodrigo, Kubo, Tanaka & Koizumi, 2014) has recently found evidence of effects of priming of word order that take place independently, and at the same time as structural priming on grammatical function assignment. In our study, we had Kaqchikel speakers[15] describe pictures after listening to descriptions made by a confederate in

[14] In Dutch, auxiliaries and participles can be used in both auxiliary-participle order and vice versa, with no change in meaning, as can be seen in the following example (from Hartsuiker and Westenberg, 2000):

(2)

(a) De man belde de politie omdat zijn portemonnee was gestolen/The man called the police, because his wallet was stolen.

(b) De man belde de politie omdat zijn portemonnee gestolen was/The man called the police, because his wallet stolen was.

Both structures are equally common in spoken language.

[15] Kaqchikel is an ergative language of the Mayan family. Its canonical word order is VOS, although the preferred word order is SVO. This language allows for free word order. Both active and passive sentences are allowed in VOS and SVO word orders (see Kubo, 2016, for more details).

SVO/VOS word order and in active/passive voice. Results showed an increase of passive sentences after passive primes, and an increase of VOS sentences after VOS primes, which suggests that the two steps involved in grammatical encoding can be independently affected by accessible information.

However, the effect of structural priming is not completely oblivious to meaning, as the so-called "lexical boost" effect suggests (Pickering & Branigan, 1998). Lexical boost occurs when the same verb is repeated in both prime (the sentence participants hear) and target (the sentence participants produce), which causes a strong increase of the structural priming effect. Mahowald, *et al.* (2016) point out that in some cases, the lexical boost effect overtakes that of pure structural priming. It seems, thus, that there is a relation between lexical information and the selected structure. However, the effect of lexical overlap between prime and target seems to be short-lived in comparison with that of structural overlap, which remains stable even with long lags of intervening sentences (Hartsuiker, *et al.*, 2008; Bernolet, *et al.*, 2016).[16]

In conclusion, speech produced by speakers of a language seems to be influenced not only by various semantic (conceptual, inherent or external to selected lemmas) factors, but also by syntactic (canonicity, syntactic movements, etc.) features. This process takes place in a highly complex interplay between lexical information (see left side of Figure 2.1 above), conceptual information (from the stage of message encoding), and implicit activation of syntactic features, jointly affecting the final shape of the utterance produced by speakers.

Note that in the study by Tanaka *et al.* (2009) mentioned above, the authors found an effect of structural accessibility, but no influence of animacy, in contrast to Bock (1986). This difference in the effects of animacy on structural priming ought to be considered in relation to the language-dependent effects discussed in the previous section. Different languages seem to allow more or less predominance of grammatical structure over animacy. In this regard, Prat-Sala &

[16] The discrepancy between the long lasting effects of structural priming vs. the short boost coming from lexical overlap can be explained both in terms of activated representation of exactly the same node, with residual activation (Pickering & Hartsuiker, 1998) or in terms of different types of memory involved in lexical retrieval (explicit / declarative) vs. syntactic retrieval (implicit / procedural) (Chang *et al.*, 2000; Chang, 2002). We will come back to this distinction in the next chapter when we discuss the different models of bilingual sentence production.

Branigan (2000) concluded their study with the claim that during syntactic processing, alternative syntactic structures compete. From this perspective, speech production is thus affected not only by conceptual accessibility, but also by the strength of alternative syntactic structures. Overall, canonical structures tend to be more highly activated than non-canonical structures (e.g. Branigan & Feleki, 1999; Prat-Sala & Branigan, 2000; Tanaka *et al*., 2011; but see Kubo, 2016). However, the precise degree of activation varies and may also be affected by external (contextual) features (Prat-Sala & Branigan, 2000).

Moreover, as we will see in the next sections, these characteristics peculiar to each language, along with cognitive factors, can affect not only the final shape of the utterance, but also the way in which incremental planning unfolds in time and the size of the chunks of information chosen as single planning units.

Timing of sentence production: What the eyes say about language production

The studies presented so far show that accessible elements tend to affect the final utterance, as they are planned earlier, and therefore assigned a grammatical function earlier or / and an earlier position at the sentence. Moreover, cross-linguistic studies show that accessible information may affect planning in different ways depending on the grammar on hand. However, these studies did not analyze the production planning online, but rather after the utterance was produced, tracking back the possible source of variation depending on the manipulated variables. The development of eye-tracking systems that allow head movements offers a possibility of analyzing speech planning as it occurs by analyzing which elements speakers are focusing on at each particular moment.

In this study, we will make use of this technology as a tool for analyzing the timing and length of sentence planning prior to speech onset. For that reason, we will first review the main experiments that have made use of this technique, and then we will relate them to the ongoing debates and research topics.

Meyer *et al.* (1998) explored this idea of the mind-eye connection (i.e. relation between the information being processed and where the eyes are directed) by applying it to spoken word production of two coordinated objects. In their study, participants had to name two objects depicted on the screen in order from left to right, by using a coordinated structure (e.g. “scooter and hat”). They manipulated object identification easiness (by deleting 50% of the contour of objects vs. full object), which related to easiness of access of the lemma; and word frequency (high vs. low), which related to access to the phonological form of the word.

Their results showed that (1) participants looked consistently to the objects in the same order as they were going to be uttered: from left to right; (2) participants started looking at the second element before the naming onset of the first; (3) more importantly, they found that both effects of object contour and word frequency affected speech onset and duration of the gazes directed to the first item. Therefore, these results suggest that speakers look at the elements they are going to produce not only in order to identify the objects, but also to retrieve the lexical form of the words. These types of gazes have been named also as "name-related gazes", as they reveal the underlying linguistic processes (Griffin, 2004).

In 2000, Griffin & Bock explored this connection further by having participants produce a transitive sentence involving two items while monitoring their eye movements. In contrast with the Meyer *et al.* study, the items not only had to be ordered, but also grammatical functions had to be assigned to them in order to produce grammatically correct responses. In the Griffin & Bock's study, participants had to simply describe a scene at the same time they were seeing it (*extemporaneous group*). There were other three groups that allowed for comparison of the different patterns found. First, a *prepared speech* group, who had to perform the description after the picture had disappeared, allowing to compare how much preparation is needed before starting formulation in the *extemporaneous group*. There were also two non-linguistic groups: *patient-detection group* (participants had to look for the patient of the action), intended to understand how fast speakers could understand the action; and an *inspection group*, which merely looked at the picture without any special purpose, and was intended to control for the saliency of one of the objects over the other. Their results showed, as can be seen in Figure 2.2, that speakers focused the subject of the sentence they wanted to construct before speech onset. Only after this moment, at the time the subject was being articulated, gazes shifted to the object of the sentence, the second uttered element. This pattern suggest a highly incremental process, in which assignment of grammatical function of the elements and lexical retrieval in the order of articulation take place simultaneously, or at least interwoven, overlapping in time.

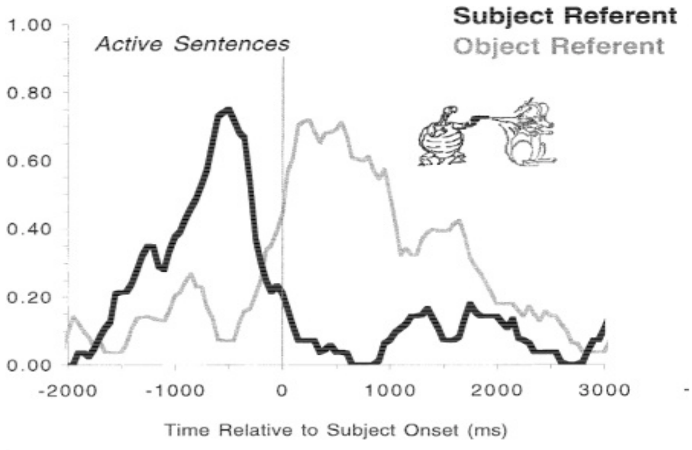


Figure 2.2. Pattern of gazes to agent / subject ("turtle") in black line and to patient / object ("mouse") in grey line when preparing and producing a transitive sentence. Time point 0 (vertical line) signals speech onset. From Griffin & Bock (2000).

However, additionally, Griffin & Bock found that prior to focusing each element in order, there was a period between 300 and 400 ms. in which speakers did not fixate preferentially either of the elements. Correspondingly, this was the period that the *participant-detection group* took to find the participant. They suggested that this period corresponded to a global apprehension of the scene, which guided subsequent gazes.

On the other hand, several studies have pointed to the idea that a global apprehension is not needed, as participants focus on the elements in order of speech without focusing nor being affected by the characteristics of the following elements. Studies like the one reported above of Meyer *et al.* (1998) point to a highly incremental process, with only one word at a time being planned. In the same direction points another study by Griffin in 2001. In this study, she had participants describe a visual array of three elements with sentences like "The clock and the television are above the needle". She manipulated the codability and frequency of the second and third nouns. Results pointed to no effects of either the codability or the frequency of any of the elements on speech latency or gaze duration to the first element, thus pointing to a highly incremental process that can be set off with only one single element. A study by Brown-Schimdt & Konopka (2008) points in the same direction. In this study, they compared the planning process of noun phrases that consisted of a noun and an adjective (i.e. "the small butterfly) in English and Spanish: in Spanish, the modifier is uttered after the noun (e.g. "la mariposa pequeña"). Their main aim was to analyze the lower level for incremental planning, so as to test whether or not planning in Spanish entails later planning of the contrasting word (i.e. whether planning at a lower level was phrasal or lexical). Their results showed that planning of contrastive elements that denote size takes place later in Spanish than in English, thus indicating that the planning unit can be as small as a single word inside a phrase. Similar results were found in Brown-Schimdt & Tanenhaus (2006) with English monolingual speakers only, in which timing of gazes to the picture with the contrasting size correlated with the position of the adjective in the adjectival phrase: either prenominal, or post-nominal due to repair.

However, these studies have all focused on small units that requires scarce, if any, syntactic processing. As Allum and Wheeldon (2007) have pointed out, the studies of Meyer and Griffin consist of repetitive tasks, with no fillers, in which participants could be relying on a strategy tailored to the nature of the task, rather than on a generalized form of speech planning. The same concern can be applied to Brown-Schimdt & Tanenhaus (2006) and Brown-Schimdt & Konopka (2008), since the task consisted of repeatedly naming a noun in contrast with a different one, without a context or a general action to be described.

Gleitman, January, Nappa & Trueswell (2007) purport to provide evidence for one of the strongest claims of lexically guided planning. In their study, they presented participants with different types of scenes (to be described by means of conjoined-NPs, active-passive predicates, symmetrical predicates and perspective predicates), and asked them to describe them. Critically, they controlled saliency by means of marking either participant involved in the scene with a visual cue of which participants were unaware, prior to picture onset[17]. They found that when

[17] In contrast with Tomlin (1997), who similarly explored the influence of visual saliency over speech, but did so with an overt cue: an arrow pointing at one the participants involved in the action that subjects had to look at during the whole task. Thus, despite the strong

the cue was effective in focusing gazes to the signaled element (which took place in a high proportion of the cases), participants were more likely to start their utterance with the cued element. Thus, they provided evidence that visual salience could guide planning, with a planning made word by word even in syntactically complex utterances: in this case, participants seemed to start articulation after the preparation of just one word, without a global apprehension of the scene.

The idea of accessibility of individual items guiding speech planning relates closely to the studies we presented in the previous section of this chapter. As was mentioned then, these studies have shown that accessible items are placed earlier or are assigned a more predominant grammatical function in the sentence. Thus, as Bock *et al.* (2004) point out, the idea of starting planning with the more accessible element in a word by word fashion is quite appealing. However, data with eye-tracking seems to point to a more complex scenario.

There is accumulating evidence showing that speech planning can take place not only word by word, in a bottom-up fashion, but also in a top-down fashion (Bock *et al.,* 2004), with the creation of a previous plan that will guide subsequent gazes. This is the result that Bock, Irwin, Davidson & Levelt (2003) and Kuchinsky, Bock & Irwin (2011) found. These studies explored how prior experience and linguistic preferences influenced the planning process when reading the time in both analog and digital clocks. Time-reading allows to explore the influence of preferred structures over initial planning (Bock *et al.*, 2003) and whether this preference can be shifted with instructions, thereby influencing the planning process (Kuchinsky *et al*., 2011).

Bock *et al.* (2003) explored latencies and fixations in time-telling with digital and analog clocks when reading time in so-called “absolute“ and “relative“ ways (e.g. “two forty” and “twenty to three”, respectively) with the main aim of comparing the results found in Griffin & Bock (2000) with the planning of a very well practiced and repeatedly uttered expression.

Their results pointed to a prior apprehension that lasted around 300 ms., and with later gazes that were in concordance with the type of expression that was

---

effects of cue found in this study, it is difficult to extrapolate its results to more natural contexts.

going to be produced: gazes first to the hour region and later to the minute region for the absolute time-telling, and vice versa for relative time-telling. These results remained constant both when participants were free to choose the structure they preferred (experiment 1 and 2), and when they were told to use one or the other switching at mid-experiment (experiment 3). Authors suggested that this results point to a process that is highly systematic, with a prior planning that led to proper gazes to those parts of the display that were relevant for the task on hand.

Kuchinsky *et al.* (2011) went further into exploring the mechanisms that control for fairly well practiced and automatized speech (as is the case for time-telling). In this case, authors asked the participants to reverse the reading pattern: that is, in the analog watch, the big hand meant the hour and the small hand the minutes. Their results showed that participants were able to shift their preferred mappings: when participants read the time in the usual way, they focused earlier in the small hand, to shift later to the big hand. However, when instructions were given to read the time in a reversed fashion, participants looked first at the big hand and then to the small one, in a very similar but reversed pattern. Thus, as authors pointed out, these results support a top-down process, in which speakers are not controlled by automatic visual processes (due to experience in always reading the small hand first), but they were able to change the plan that guided their gazes, in this case, in an unfamiliar way.

Evidence for top-down processes is not only supplied by studies on time-telling, but also in research exploring the production of transitive sentences (with content and actions speakers have probably not seen before, as Bock *et al.* (2003) pointed out). In this regard, Ganushchak, Konopka & Chen (2014) analyzed eye-movement patterns in Dutch and Chinese speakers in response to questions about a picture participants had to describe. Prior questions were either neutral, focusing on the agent or focusing on the patient, as can be seen in (12). Both languages exhibit a contrast in the way questions are formulated: Chinese wh-questions are formulated in situ, that is, they lack overt movement as in Dutch. This difference leads to differences in the way in which patient-focused questions are formulated. Specifically, participants in Dutch have to elaborate a proper sentence frame to answer the question in this case, while for Chinese speakers they only have to repeat the frame they have just seen with the new, focused element.

(12)

Targeted utterance: "The policeman is stopping the truck"

Neutral question: "What is happening here?"

Agent-focused question: "Who is stopping the truck?" (Dutch and Chinese)

Patient-focused: "What is stopping the policeman?" (Dutch) "The policeman is stopping what?" (Chinese)

After presentation of these questions, participants had to answer while looking at a picture with a transitive action on it. Similarly to Bock *et al.*'s and Kuchinsky *et al.*'s results, these authors found that within 400 ms. speakers rapidly directed their gazes at the new element (the element that was not provided in the question, and needed to be coded), with questions modulating gaze patterns. Interestingly, Chinese speakers started focusing on the object earlier than Dutch speakers in the Object-focused condition (400 ms. vs. 800 ms.), suggesting that there was no need of a previous structure encoding before accessing to the next lexical item (i.e. the object), in contrast with Dutch, in which the structure to be created differed from the one presented in the question. Thus, these results suggest that prior planning of the relation of the elements was needed before participants started linguistic encoding. In Chinese, since this structure was already provided, planning could move rapidly towards the encoding of the new element. It also shows that linguistic encoding was efficient with gazes located at the elements that were not provided and, therefore, needed further lexical codification.

Interestingly, although the cross-linguistic differences found in the previous study were not striking, they point to differences in the way speakers of different languages undertake speech planning, depending on specific patterns of each grammar. Thus, increasing evidence making use of eye-tracking persuasively suggests that whether or not there is advanced planning of the upcoming structures seems to depend on several factors, either within-language, such as the accessibility of the lexical items or structural relations or event information (e.g. van de Velde, Meyer & Konopka, 2014; Konopka, 2012) or between-languages, such as basic word order in canonical sentences (Hwang & Kaiser, 2014, for Korean; Norcliffe *et al.*, 2015, for Tzeltal; Kubo, 2016, for Kaqchikel; or Sauppe, Norcliffe, Konopka, Van Valin & Levinson, 2013, for Tagalog) or in specific structures (*wh*-questions, as we saw before in Ganushchak *et al*., 2014).

Recent studies by Norcliffe *et al.* (2015) and Kubo (2016) attempt to analyze the effect of saliency (in terms of animacy) on participants' gazes and utterances when producing sentences in a language whose canonical order highly differs from English: Norcliffe *et al.* targeted Tzeltal, while Kubo examined Kaqchikel, both mayan languages with a VOS canonical word order. These languages accept a SVO word order as well.

Kubo (2016) found that when the agent was animate, participants were more likely to produce VOS sentences than when it was inanimate. Similarly, both studies found that speakers start focusing on the agent early in the planning course, despite its being placed at final position. These authors concluded that lexical competence (see the Gennari *et al.* study cited above) could be involved in the choice of word order, preferring SVO orders when elements were more similar, thus separating them in time. However, a certain degree of lexical competence, even if coming from a conceptual, rather than grammatical (of lemma access) level, implies a certain degree of global planning of both agent and patient, in order to decide which structure is more appropriate to use. Moreover, Kubo also found that, in VOS sentences, participants tended to fixate the agent of the scene until around 1000 ms. After that, speakers started fixating the patient before speech onset. Remember that the first word uttered here is not the patient, but the verb. Thus, it is likely that participants are elaborating the structural relations among the different elements by first focusing on the subject, which is the syntactically most prominent element, and then devoting time to building the grammatical relations with the patient, in order to articulate the verb.

A remarkably similar pattern was found in Tzeltal (Norcliffe *et al*., 2015). In this case, the authors compared the gaze patterns while producing transitive sentences in both active and passive voices in Tzeltal (with SVO and VOS orders) and in Dutch (with only an SVO order). Importantly, gaze patterns in Dutch and Tzeltal SVO order were strikingly similar, with participants encoding the elements in the same word order as they were going to be uttered. However, similarly to the Kubo study, when a VOS is produced, the gaze pattern shows that speakers prioritize structural information over linear information. In both Norcliffe *et al.*'s and Kubo's studies, previous appraisal of the animacy of both agent and patient led speakers to choose SVO or VOS orders. Thus, as Norcliffe *et al.* point out, even in languages with SVO order there is a certain degree of high level processing involving agent, patient and their relations. This processing guides subsequent gazes to the different elements, and largely determines the selected word order.

It should be noted that both Kaqchikel and Tzeltal exhibit a very complex verbal morphology, in which number and person morphemes of both agent and patient must be appended to the verb in order to articulate it. Although this might introduce a bias toward an early encoding of both agent and patient, there is also evidence of early encoding of the global scene in another verb-initial language (namely Tagalog), where both subject and object (instead of the verb, which marks only voice and tense) are marked with grammatical case marker before each noun (Sauppe *et al.,* 2013). Moreover, Konopka & Kuchinsky (2015) analyzed the effect of event accessibility and structural accessibility on the time-course of sentence formulation, in Dutch verb-initial and verb medial-sentences. Their results showed a stronger effect of event coding in early time windows in verb-initial constructions than in utterances with SVO word orders. Thus, taken together, these studies point to a flexible pattern of sentence planning, which can be adapted to the particular characteristics of each language. Importantly, however, they not only show how languages may differ from one another, but also how languages that are widely different (as Dutch and Tzeltal) can share the same planning processes and timings when the produced utterances are similar. Additionally, they underscore the fact that even within the same language, variations in word order lead to different planning processes, showing that there is flexibility also within languages.

In summary, all the studies with eye-tracking methodology reviewed so far reveal a fairly complex reality. Speakers are able to start speaking after the preparation of a single element, but they also show a tendency to create a global encoding of the scene that will guide encoding. This seemingly contradictory evidence has led to two different positions regarding sentence planning and the way it proceeds incrementally:

- *Linear or Lexical Incrementality accounts:* This position posits that production planning can start with the minimum information needed at any time. On this account, characteristics of single lexical items, such as their accessibility or the ease to name them, make them more likely to be selected as the starting point of the sentence, thus engaging the formulator in the preparation of this single element. From this perspective, grammatical relations are encoded after the first element is decided, as a consequence of this selection. This account is based on the evidence that shows a tendency of speakers to place the animate or more salient element at the beginning of the sentence (e.g. Bock & Warren, 1985; McDonald *et al.*, 1993), as on the evidence showing that visual cues correlate with starting points (Gleitman *et al.*, 2007; Tomlin, 1997) and that under certain circumstances speakers start articulation after fixating and preparing only the first lexical item (Griffin, 2001; Meyer, *et al.,* 1998; Brown-Schimdt & Konopka, 2008; Brown-Schimdt & Tanenhaus, 2006).
- *Hierarchical or Structural Incrementality accounts:* This perspective, on the other hand, states that production planning is guided by an early gist of the event to be described. This gist includes conceptual characteristics of the event, as well as the structural relations between the participants involved. The creation of this structure will subsequently guide the retrieval of lexical elements in a top-down fashion. Thus, contrary to the previous position, structural incrementality posits that relational processes (i.e. the relation between elements involved in the to-be-uttered scene) play an earlier and more definite role in utterance planning. Evidence for this position comes from studies in which participants had to describe a scene freely (Griffin & Bock, 2000) or under conceptual or grammatical constraints (Van del Velde *et al.*, 2014; Ganushchak *et al.*, 2014). Similarly, time-telling studies (Kuchinsky & Bock, 2011, Bock et al, 2003) and cross-linguistic evidence (Hwang & Kaiser, 2014; Norcliffe *et al.*, 2015; Kubo, 2015; Sauppe *et al.*, 2013) seem to support this position (see Bock *et al.,* 2004 and Norcliffe & Konopka, 2015, for a review of both perspectives).

However, as a matter of fact, these two positions stand more as two extremes of a continuum, rather than clear-cut positions. Similarly to the issue of speech planning scope, which will be reviewed in the next section, research on eye-tracking points to a more complex reality. Flexibility in speech-planning, both within and between languages, shown by prior studies implies that speakers can create a global structure, that guides the encoding of each of the elements, rather that the other way around. However, what this evidence also shows is that under certain circumstances, a more lexically-driven planning is also possible. Thus, languages like English allow participants a greater degree of linear incrementality (as seen in Gleitman *et al*., 2007). This language is subject initial, allows little word order variation, and does not have case marking of subject or object. Similarly, syntactically simple sentences (Griffin, 2001; Meyer *et al*., 1998) or phrase-long utterances (Brown-Schidmt & Konopka, 2008) can lead to automatize planning processes that will allow lexically-driven planning (see Norcliffe & Konopka, 2015, for a review).

Different reasons can lead speakers to rely on either way of planning, increasing or reducing the scope of planning. In the next section, we will introduce some studies that have explored the length of speech planning scope, as well as its flexibility.

Planning scope: Length of speech planning units

Another important topic in sentence production that has received widespread attention, and that is closely related to our research, is the size of the chunks of information that speakers plan at any given time. A distinctive feature of incremental production is the fact that speakers do not wait until the whole utterance is prepared at one level of processing, before starting processing at the next level (Kempen & Hoenkamp, 1987; Levelt, 1989). However, at the same time, speakers do not process the message word by word, or phoneme by phoneme. Understanding the length of the “chunks of information” that speakers handle at any step of the planning process is vital to understand what kind of information determines the final utterance, and how it is used along the process (i.e. how accessible elements are used during the different steps of speech planning). Although planning scope is not the main focus of our research, it is necessary to introduce here briefly the main findings in this field, in order to understand how a speaker can balance between structural and conceptual information (regardless, or in addition to, accessibility). A preliminary remark concerning this issue is that most studies on planning scope are limited to simple clauses or sentences with relatively low complexity, so there is not much reliable information as yet about the upper limits of planning scope in production research.

Planning scope research dates back to speech errors studies (Garrett, 1980). Among other things, speech errors have been an important source of information about the different planning units that speakers handle at any point of the production process. Although production usually flows in a fluent fashion, without significant interruptions or errors, and thus allowing for efficient communication, this is not always true, and speakers sometimes make mistakes when speaking. These errors, which are part of normal speech, are highly systematic and can be classified. Different categories of errors depend on whether the source of the error is present or not, and the type of "disturbance" that occurs (Igoa & García-Albea, 1999). One broad category, involving errors in which the origin of the error is present, is also known as "contextual" errors. Especially this type of errors has allowed estimating the extent of planning scope at different levels of processing. They show that different elements of the speech are planned together, before being misplaced in a different position than was intended. Different patterns that can be observed through their analysis are that (examples from Fromkin, 1973):

1. Errors that involve words (e.g. exchange of two words in the utterance), usually exceed the phrase boundary, moving inside the *same clause*, and occasionally between adjacent clauses. E.g.

   (13) Word reversal: "the *cleaning* of the *cost* of the carpet" (intended: "the cost of the cleaning of the carpet").

2. Errors that involve word stems often take place within a single phrase or span two adjacent phrases within a maximal projection.

   (14) Stranding of grammatical morphemes: "A hole full of floor*s*" (intended: "a floor full of hole*s*").

3. Errors that involve *phonemes* almost always occur within the boundaries of a single *phrase*, even inside a content word or across two contiguous words.

(15) Single sound anticipation: “Ki*n*chen sink” (intended: “kitchen sink”).

These different observations have led to empirical research whose main aim was to clarify the length of sentence planning units prior to speech onset, and whether sentence planning is conceptually or grammatically guided (i.e, lexical or syntactical). Error analysis points at a scope that is often equivocal between a clausal level, a phrasal level and even a word level, and probably depending on planning level (i.e. conceptual, grammatical, phonological) that is being carried out at each particular stage. However, the observational nature of speech error methodology makes it difficult to test hypotheses regarding this and other empirical issues in language production in a reliable way. Fortunately, there are several experimental methods in the production literature that may help clarify the planning scope issue, such as subject-verb agreement production studies (e.g. Bock & Miller, 1991; Bock & Cutting, 1992; Vigliocco, Butterworth & Garrett, 1996; Haskell & MacDonald, 2005; Veenstra, Acheson, Bock & Meyer, 2014), semantic interference procedures through a distractor word in verb planning (e.g. Schriefers, Teruel & Meinshausen, 1998; Momma, Slevc & Philips, 2016; Hwang & Kaiser, 2014), or the effects of NP complexity on speech onset latency (Smith & Wheeldon, 1999; Allum & Wheeldon, 2007; Wheeldon, Ohlson, Ashby & Gator, 2013), sometimes paired with eye-tracking (Meyer *et al*., 1998; Griffin, 2001; 2003).

The studies reported below show mixed evidence pointing to different length of scope: they may be divided between those that point to a phrasal planning scope, and those pointing towards a clausal scope. But more importantly, it seems that these discrepancies are due to the level of planning that is being assessed in each of these studies. Thus, evidence pointing towards clausal scope can be found, for instance, in Meyer (1996), who found semantic interference on lexical items that appear later in the utterance, thus showing some degree of grammatical planning (i.e., access to the lemmas). However, as Meyer acknowledges, these semantic interference effects might be purely conceptual. On the other hand, as was discussed in a previous section, Griffin (2001) and Meyer *et al.* (1998) also found evidence for a smaller scope, with participants not being affected by distractors that are related to the second or the third NP. Schriefers & Teruel (1999) and Griffin (2003) analyzed the access to phonological form when planning two word utterances. They found that, when the first word is long enough to ensure fluent speech, speakers do not prepare the phonological form of the second word before speech onset. These studies show a picture of a progressively limited scope in language production as the system engages in more atomic parts of the utterance: from a more global, clausal level, based on conceptual planning, to a word by word, or even syllable by syllable (Schriefers & Teruel, 1999) planning. These studies also show the flexibility of the system, and its ability to determine whether or not the first word or the first phrase is long enough to ensure fluency without preparing the second one (Griffin, 2003).

High-level (conceptual) planning and lexical or syntactic planning must necessarily interact in sentence production, with at least some parts of both having to be prepared before speech onset. To explore how much of either type of planning takes place before speech onset was the aim of Smith & Wheeldon's (1999) study. In their experiments, they presented participants with a set of three pictures that were performing motion. Participants had to describe them with either of the following utterances (from Smith & Wheeldon, 1999), labelled as follows:

(13) Complex-simple sentence: The dog and the foot move above the kite

(14) Simple-complex sentence: The dog moves above the foot and the kite

(15) Complex sentence: The dog and the foot move up

(16) Simple sentences: The dog moves up

Results showed that speech onset latency was larger in (13) and (15) than in (14) and (16) respectively, thus showing that in the latter case only "the dog" is being prepared prior to onset, while in sentences with an initial complex phrase the whole "the dog and the foot" is under preparation. Subsequently, in order to explore whether the effects were due to lexical access or syntactic planning, they repeated the same task adding a preview of the pictures beforehand. Results showed a significant decrease of speech onset in all conditions, showing that lexical access accounts for the bulk of time devoted to preparing the utterance. Still, differences were found mainly between the simple (16) and complex sentences (15), which shows there is some syntactic planning involved in the preparation of the first complex phrase. Importantly, the preview of the lexical items had more impact in reducing the onset of sentences starting with complex clauses than of those starting with simple clauses, suggesting that lemma preparation spans the entire clause. These authors also noted that being exposed to the first noun had a greater impact in reducing speech onset than did being exposed to the second one, thus suggesting that the processing of the first part is done more thorough than the second one. Allum & Wheeldon (2007) refer to this scope as the first verb argument phrase[18] in speech planning. In English, both the initial element and the whole verb argument phrase play a key in the preparation of the grammatical function (subject in this case). However, in head final languages, the first element and the main element of the phrase do not always match. To explore this, Allum & Wheeldon (2007) used the same task as Smith & Wheeldon, but with Japanese as the target language. As can be seen in (17) to (19) (extracted from Allum & Wheeldon, 2007), the first element of the complex subject phrase (Inu-no ue-no; "above the dog") is neither dominant nor vital for the production of the subject in the case of Japanese, in contrast with English.

(17) [Inu no ue no] hana wa aka desu.

Dog-GEN above-GEN flower-TOP red is

"The flower above the dog is red."

(18) [Inu no ue no] [hana to zubon to ringo wa] aka desu.

Dog-GEN above-GEN flower and trousers and apple-TOP red are

"The flower and apple and trousers above the dog are red."

(19) [Inu to hana no ue no] [zubon to ringo wa] aka desu.

Dog and flower-GEN above-GEN trousers and apple-TOP red are

"The trousers and apple above the dog and flower are red."

[18] Verb argument phrase is a major unit in the grammatical structure of a sentence, subject or object, and all its related arguments. For example, in the sentence "The bird above the tree is a hawk", the first verb argument phrase is "The bird above the tree" (Allum & Wheeldon, 2007).

Their results replicated those found in Smith & Wheeldon with a head-final language. The first phrase, before the genitive, was more thoroughly planned than the subsequent elements, with different latencies depending on its length. However, planning spans the whole clause, with high-level planning occurring beforehand.

Similarly, the role of the verb in language planning has been extensively discussed, with many influential models of language production bestowing it a central role in planning (e.g. Bock & Levelt, 1994; Kempen & Hoenkamp, 1987; Levelt, 1989). This debate addresses the issue of to what extent and how thoroughly is the verb planned before speech onset, thereby revealing the length of planning scope and whether it varies cross-linguistically. One rather extreme approach to this issue claims that the verb is selected early in sentence planning, even before the phonological encoding of the first element has taken place (Ferreira, 2000). Despite this emphasis on the role of the verb as a crucial piece of information in planning, the evidence is not clear, especially when taking into account cross-linguistic differences in word order. Kempen & Huijbers (1983) show evidence for an advanced verb selection: they found that changing the verb between trials had an effect on both SV and VS sentences, positing that in both cases the verb was accessed prior to speech onset. However, Schriefers *et al.* (1998) ran a similar experiment with a semantic interference task. In this case, they compared the speech onset latencies in SV and VS sentences and in SVO and SOV sentences in German. Their results showed that there was no semantic interference in either the SV or the SOV groups, but there was in verb initial utterances. These authors concluded that prior access to verbal syntactic information is not necessary in order to articulate an utterance. Thus, conceptual information of the scene and the relation between its elements can be enough to guide the planning process and start speaking. However, when verb information is available, speakers also make use of it to plan their speech. Thus, these authors infer that there is a dual route to sentence planning, conceptual and lexical, and that speakers move from one to the other depending on the information available during planning. Momma *et al.* (2015) provide further evidence in support of this claim, by analyzing verb advance planning in Japanese, a head final language that allows considerable flexibility in constituent order (with the only restriction that the verb must be placed at the end of the sentence) and NP dropping. They compared advance planning of SV and OV sentences in which either the object or the subject, respectively, was dropped. Results showed advanced planning only in OV sentences, but not in SV sentences. Momma *et al.* reasoned that verb advance planning takes place only before planning the object, because this constituent is more dependent on verb. The characteristics of the verb, and the arguments it takes, determine the features of the object or even its sole presence of absence. This dependency is not present in subject-NPs, which are at the same high level than verb phrases in the syntactic tree. Thus, this study points to the fact that whole clauses, centered on the head, are planned as a whole in verb-final languages, despite final position of the verb in the sentence. However, what these authors could not disentangle with their results, as they recognize, is whether they are due to the conceptual planning of the message or to the grammatical encoding of the verb and its properties.

Nevertheless, there is further evidence that underscores the idea that syntactically dependent elements are planned together beforehand. Lee, Brown-Schmidt & Watson (2013) explored the planning scope of relative clauses to see whether advanced planning took place linearly (i.e. encompassing elements that are close to each other in the utterance) or hierarchically (i.e. encompassing elements that are closely related syntactically, even if they are far from each other in the utterance). They made participants produce high and low attachment relative clauses, in which the HN inside the subordinate clause was either medium or highly codable. In high attachment relative clauses (e.g. "The apple of the student that is on season") the relative clause ("that is on season") is directly dependent on the first noun ("the apple"), while in low attachment relative clauses (e.g. "The apple of the student who is in class"), the relative clause is only indirectly dependent on the first HN ("the apple"), while directly depends on the second one ("the student"). Their aim was to explore the extent of advanced planning in relative clauses and their syntactic relations. Their results show that duration of the first noun was longer when the third noun (the subordinate one) had medium codability compared to when it had high codability. Importantly, this effect was found equally in both high-attachment and low-attachment relative clauses, showing that in both cases the relative clause was planned along with the articulation of the first noun. This study suggests that, even if dependent constituents are long and spatially separated, hierarchical relations are planned together prior to the utterance, thus guiding subsequent lexical access.

Taken together, these results provide a picture in which planning scope is flexible and can vary from one single word to a whole complex clause, moving from linear to hierarchical incremental planning. The specific size of planning units is affected by multiple factors, from task demands (Wagner, Jescheniak & Schriefers, 2010) to time pressure (Ferreira & Swets, 2002) to working memory capacity (Swets, Jacovina, & Gerrig, 2008; Slevc, 2011). Konopka (2012) analyzed the effect of structural and linear information over planning scope, by controlling structural accessibility and lexical accessibility. She observed that early easy to retrieve lexical items resulted in an enhanced planning scope, with speakers starting to plan the second element before speech onset. However, this effect was only present when structure was accessible (i.e. when the structure was previously primed). In general, accessible structures produced shorter onset latencies, regardless of lexical accessibility. Thus, Konopka concluded that, although planning scope is highly sensitive to cognitive and linguistic demands, the default planning process is one that starts with the structural relation of elements. Once structural relations are mastered, lexical elements come into play.

Summary and conclusions

In this chapter, we have presented the main lines of research that have attempted to understand the way humans produce language. In general, the picture that can be draw from all the research addressed is that of a mechanism in which there are several stages involved in transforming the message (a global conceptual representation) into a string of articulable sounds that abides by the grammar of a given language and correctly expresses the message to be conveyed. In order to do so, the system gradually handles more atomic chunks of information, from clausal to phrasal level, and to subphrasal level, in an incremental fashion, that is, without waiting until all the information at each level is available, but working instead with small chunks of information as they are prepared. At each level, different operations are executed: thematic role assignment, grammatical role assignment, lemma selection, linear assembly of elements, phonological encoding, and so on.

This complex system is characterized by an interplay of several linguistic (e.g., grammatical features of the spoken language, word frequency) and cognitive (e.g., conceptually prototypical representation, overload of cognitive resources, conscious planning strategies) factors, giving rise to a picture that is difficult to fully apprehend. All these variables provide an idea of the speech planning system as a highly flexible apparatus in which structural relations (relational information) and lexical information (non-relational information) converge (Konopka & Meyer, 2014).

Conceptually salient information during message encoding exerts an effect on linguistic encoding in either grammatical function assignment or easier lexical retrieval (placing elements first). Similarly, readily available structural frames (either by remaining activation or by learning through exposure) influence the final shape of the utterance: the role that each element is going to play, the order and the relations that they are going to take. It seems clear that language production consists of a continuous interplay between relational and non-relational information. However, there remain to be explored the role that each of these two types of information play along the incremental process of speech planning.

1. Results from studies assessing the effects of conceptual accessibility and eye-tracking studies, and from those exploring the time-course of sentence formulation alike have difficulties in distinguishing whether the observed effects are coming from relational or non-relational information: the first element also happens to be the subject. Thus, it is not clear whether early planning relies more strongly on the relation between elements or on the activation of single items. First of all, the role that conceptual accessibility of animate elements plays in the starting point of speech planning, that is, in selecting where and how the utterance should start, is still not clear. In the second place, eye-tracking studies have shown cross-linguistic differences in planning. Unfortunately, all these sentences had the most dominant element at the beginning, usually at subject position: gazes are directed to this element right after a short 400 ms. interval. However, these gazes do not inform us of whether structural planning is taking place (as the subject is the most dominant element in the sentence), or if they are devoted to retrieving the name of that item to be placed in order.

Therefore, a study aiming to test a structure that allows to disentangle both is highly advisable.

2. As Bock (1986) showed in her study, conceptual accessibility and structural accessibility both play a role in determining the final shape of the utterance. Relational and non-relational information play a role, not only in choosing the beginning, but also along the whole sentence formulation process. However, the time-course and the relative weight given to conceptual accessibility and to structural accessibility are not clear.

In this PhD Thesis, we will further explore the interplay between relational and non-relational information by presenting a study in which the effects of both kinds of information can be fully disentangled in speech planning. As it was presented in Chapter 1, RCs in Japanese, in contrast to Spanish, constitute an ideal case to explore these questions: the HN is placed at the end, after all its subordinate elements have been presented. Moreover, it is important to note that, through the exploration of a head final complex structure as RCs in Japanese, we may also examine the extent of planning scope for these sentences, as we analyze what information speakers take into account at each moment of the planning process. The comparison with a head initial language, like Spanish, would then allow to compare the role of linear incrementality and hierarchical incrementality from a cross-linguistic perspective.

Additionally, the comparison between RCs with the patient as HN in active and passive voice in both Japanese and Spanish would allow, to deepen our understanding concerning, not only what information is prioritized when linguistic encoding begins, but also how relational and non-relational information become intertwined with each other along the whole process, even after speech onset has started. This is possible thanks to the fact that word order remains stable between active and passive sentences within these two languages.

Nonetheless, before presenting our studies on this topic, let us introduce a brief review of the literature available so far on bilingual sentence production, in order to set the ground for the third experiment of this research project, meant to explore the planning system in bilingual speakers of two typologically distant languages: Spanish and Japanese.

Chapter 3. Sentence Processing and Production by Bilingual Speakers and Second language acquisition processes

The field of sentence production by bilingual speakers is a somewhat underdeveloped one in psycholinguisitc research, with many unresolved questions. Nevertheless, some authors have tried, from different perspectives, to explain how people with more than one language produce sentences in their second language and what mechanisms are involved. In this chapter, we will first review the main models of bilingual sentence production. After that, we will examine some studies on bilingual sentence comprehension and production with special emphasis in the incrementality of those processes, and the role that accessibility plays on them.

Models of bilingual production

The main focus of bilingual sentence production studies has been to clarify whether the rules and processing strategies from a bilingual's L1 and L2 are separately represented and/or have some overlapping points (Sanz & Igoa, 2012). In general, they show evidence of shared activation of syntactic features in bilingual speakers, with some differences in the processing patterns of bilingual vs. monolingual speakers. This evidence has guided both early and current models of bilingual sentence production, one of whose major challenges is to account for syntactic coactivation while at the same time selecting and speaking in one of the languages, and keeping the other under inhibitory control. To our knowledge, there have been only a few attempts to model the speech production system within a bilingual mind.

The first attempt to explain planning processes in bilinguals was a proposal by De Bot (1992), who adapted Levelt's model of monolingual language production to bilingual speakers. In the first place, this model suggests a revision of Levelt's idea of the conceptualizer as language specific, by suggesting that there are two separate parts in the bilingual speaker's production apparatus, namely, the macro-planning component, which is not specific to a particular language, and the micro-planning part, which is sensitive to the requirements of each of the languages that a bilingual speaker commands. In the second place, De Bot assumes that there are two formulators in the bilingual's mind, which are, otherwise, connected. The degree of interaction, or closeness, between formulators varies depending on language distance (closer languages having closer formulators) and proficiency (high-proficiency bilinguals having more separate formulators). Remarkably, there is only one mental lexicon that is shared by both languages of the bilingual speaker. Lemmas of different languages are distributed in subsets, which are selected through inhibitory control (following Green's inhibitory control model (e.g. Green, 1998)). Finally, the message plan created for the formulator of the selected language is sent to an articulator that is shared between languages. Figure 3.1 shows the representation that Hartsuiker & Pickering (2008) presents of De Bot's bilingual production model.

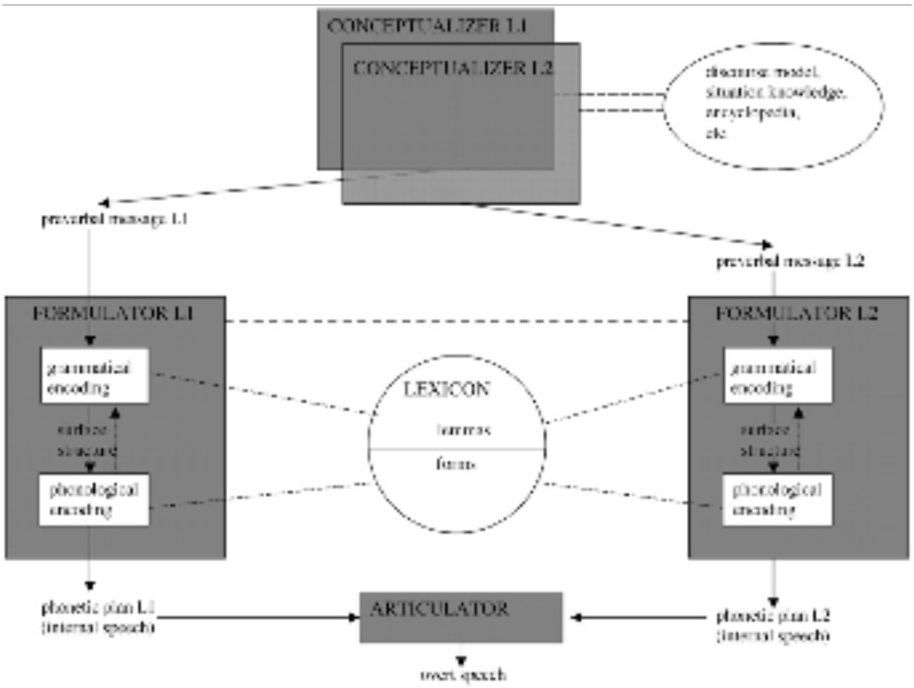


Figure 3.1. Representation of De Bot's model of bilingual production (1992) (from Hartsuiker & Pickering (2008)).

A second model of bilingual sentence production is the one proposed by Hartsuiker, Pickering and Veltkamp (2004), and Hartsuiker and Pickering (2008), which they named "Integrated model of bilingual language representation". These authors, encouraged by the evidence of cross-linguistic structural priming, proposed a model that emphasized the close relation of languages at the level of grammatical encoding (the most distant component in the bilingual production system according to De Bot). They focused their proposal on the mapping between the mental lexicon and syntactic encoding, from a lexicalist approach. They claim that lexical information is established in a series of connected nodes at different strata. In particular, they consider the existence of three strata inside the lexical representation of the items:

conceptual, lemma and word-form, with syntactic information, as well as word order, being contained within the second one. Importantly, this model also considers that the mental lexicon is shared by the two languages, resulting in combinatorial nodes that are connected to all words that share similar characteristics, irrespective of language. The state of activation in the system results from activation spreading towards connected nodes regardless of the language, rather than from the inhibition of the non-target or latent language. The differential access to L1 or L2 is, then, accomplished thanks to an extra 'language node' (such that L1 lexical entries are associated to the L1 node, but not to the L2, and vice versa). The lexical boost that was presented in the previous chapter, or its equivalent in bilingual structural priming, the 'translation boost', are explained in terms of nodes whose activation is enhanced due to close representations: all features are shared, thus maximizing the spreading of activation. A representation of this proposal can be found in figure 3.2, for the lexical representation of Spanish-English bilinguals. In this model, different structures with same word order across languages are shared between them. However, when word order differs, a different node, specific for that language, is created.

According to this model, language distance does not affect connections between languages, to the extent they share nodes. A natural consequence of this for unrelated languages is that there are fewer structures that share syntactic features as well as word order. However, if this is kept constant, nodes will be shared by both languages. Although in this early version of the model, the authors proposed that proficiency did not affect the extent in which these connections were made (as long as the construction has been acquired, nodes will be created) (Hartsuiker & Pickering, 2008), later versions (e.g. Bernolet, Hartsuiker & Pickering, 2013) acknowledge that language proficiency has effects in the nature of shared syntactic representations, which has caused an extension of the model to account for these differences. Based on their results with more and less proficient bilinguals, they propose that less proficient bilinguals create language-specific representations for specific items of the L2. Bilinguals with low proficiency rely to a greater extent on lexical information, showing a larger effect of the translation boost. As bilinguals attain proficiency, these syntactic representations become more abstract and evolve to be shared with the L1, thus creating a single network of interconnected nodes. Figure 3.2 shows the representation of this model in bilinguals with less (left) and more (right) proficiency in their L2.

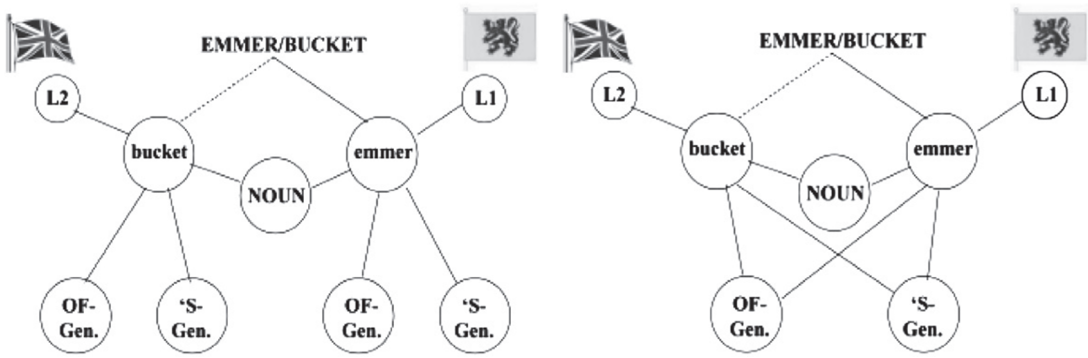


Figure 3.2. Representation of the integrated model of bilingual language representation by Bernolet *et al.* (2013): representation of the lexical entries of the noun "bucket" in Dutch and English for a Dutch-English bilingual speakers, who are less proficient (left) and more proficient (right).

Additionally, as Kantola and Van Gompel (2011) posit, those syntactic representations that are not completely equal between languages (both in terms of argument structure and word order) would likely be represented in a separate representation for each language.

Both approaches (Hartuiker *et al.* and Kantola & Van Gompel) explain the connection between languages in terms of spread activation between nodes. Accordingly, then, the effect of one language over the other is enhanced by the activation of closely related lexical elements (lexical boost, as was described in chapter 2), but these effects are not expected to last long. On the other hand, Chang *et al.* (2006), and Chang (2002) propose an *Implicit Learning Account*. This model explains the way in which humans learn language in terms of representations that are constantly changing according to the input and the match or mismatch with previous predictions (prediction error). According to this view, structural priming, both within and between languages, would be a “side effect” (Bernolet, 2008) of this mechanism. The mechanism of prediction error explains why structural accessibility is more prominent when the prime is a less frequent structure than when it is an expected utterance. It also predicts long lasting effects of the prior heard structures in the production of speakers, but does not predict effects coming from lexical overlap between elements.

Finally, Ullman (2001; 2004) focuses on a different level of representation: his proposal is centered on a neuro-anatomical stratum of language representation. This model assumes that the representations of lexical and syntactic information engage different types of memory (and are located, consequently, in different parts of the brain): declarative memory is in charge of lexical information, containing all the information in the mental lexicon[19], whilst sentence processing (syntax rules) engages the procedural memory system[20]. In this proposal, late bilinguals cannot rely on their procedural memory when using the L2 grammar to the same extent as monolingual speakers do, thus having to shift to their declarative memory when carrying out tasks that involve syntactic processing. For this reason, late learners store grammatical knowledge in their lexical representation, in an overlearned, unanalyzed pattern. However, the representation that bilingual speakers hold vary with proficiency, shifting to a progressively greater use of the procedural system as they attain proficiency in their L2. Thus, this model assumes more overlap between the languages' grammatical systems (as they share location in the memory systems) at high levels of L2 proficiency.

All these models differ basically in their level of analysis, as well as the role they assign to cross-linguistic interaction and overlap. Hence, as Hartsuiker and Pickering (2008) and Hartsuiker, Beerts, Loncke, Desmet & Bernolet (2016) noted, the three models assume different predictions concerning cross-linguistic influences. They assume (in their latest versions) assume that there is an effect of proficiency. However, both Ullman and Bernolet *et al.* hypothesize that cross-linguistic influence increase as language proficiency does. This is due to the fact that more proficient bilinguals share the same network of interconnected nodes with the L1 (Bernolet *et al.*, 2013), with a shift of the L2 syntactic processing to the same procedural memory areas than L1 (Ullman, 2004). However, for Ullman, in contrast with Hartsuiker & Pickering and Bernolet *et al.*, L1 and L2 grammars are stored in two closely related but separated storage spaces, with separate representations of syntax for both languages. On the other hand, Hartsuiker and his collaborators postulate that

[19] Located in the temporal lobe and in subcortical areas of the temporo-medial lobe (specifically, the hippocampus).

[20] Located in the connections between the subcortical areas of the basal ganglia, areas of the left inferior-frontal lobe, the left superior temporal lobe (where it overlaps with declarative memory), areas of the parietal lobe, and cerebellum.

eventually, once the bilingual is competent in her L2, both L1 and L2 will be stored in the same network of interconnected nodes. These two models assume that linguistic distance does not affect the architecture of production processes as long as the grammatical structures are shared between languages, both in terms of function and in word order (Hartsuiker *et al*., 2004; Kantola & Van Gompel, 2011), since the neural areas are shared regardless of language (Ullman, 2001; 2004). On the other hand, De Bot (1992) states that the two formulators a bilingual has become more separate as language proficiency increases, since the bilingual speaker does not have to rely on the L1 to complement the L2 linguistic system. Even at low proficiency levels, both formulators constitute separate entities according to this model, although there is a connection between them. Similarly, and in contrast with the two previous models, De Bot postulates that more linguistically distant languages have more separate and, thus, less connected formulators than closer languages. Regarding the nature of the connection between languages, Hartsuiker *et al.* (2004) propose an activation spreading mechanism, in contrast with Chang *et al.* (2006). De Bot (1992), in turn, proposes that lexical selection results from inhibitory control, although he is not explicit on the relation between formulators.

These different models constitute the first steps to understand the bilingual linguistic system. They are based on an incipient amount of research that focuses on the distinctive characteristics of bilinguals, in comparison with monolingual speakers. In the next section, we will describe some of these studies, to focus afterwards on accessibility factors and incremental production.

Defining characteristics of bilingual speakers

Qualitatively different speakers?

Speakers with more than one language show idiosyncrasies that differentiate them from their monolingual counterparts. There is a debate whether these differences can be attributed to either qualitative or quantitative differences in the way they produce and understand language (Runnqvist, Fitzpatrick, Strijkers & Costa, 2012; Roberts, 2013). Production studies have focused mainly on lexical production. In order to produce words in their L2 (or even in their L1), bilingual speakers have to be able to efficiently select the lexical item in the target language, avoiding massive interference from the non-target one. In order to do so, some authors have proposed a control mechanism restricted to bilingual speakers and specifically dedicated to lexical-selection, a mechanism that is not functional in monolinguals (language selective access (e.g. Kolers, 1966; Macnamara, 1967; Meuter & Allport, 1999; Costa, Miozzo & Caramazza, 1999; Costa & Santesteban, 2004a)). On the other hand, lexical selection has been analyzed not as a qualitatively different process, but as a result of a more effortful and costly processing that is, otherwise, present in monolinguals as well (language non selective access (e.g. Beauvillain & Grainger, 1987; Poulisse & Bongaerts, 1994; Dijkstra & Van Heuven, 2002; Dijkstra, Miwa, Brummelhuis, Sappelli & Baayen, 2010; La Heij, 2005)) (see De Groot, 2011 for a review of both perspectives; and Runnqvist *et al.*, 2012 for the implications that both positions have on the qualitative-quantitative debate). Moreover, bilingual and monolingual lexical access does not differ only in terms of language control, but also in terms of decreased speed and accuracy of lexical retrieval in the case of bilinguals. This has been seen as a possible evidence for the inhibition account, which postulates qualitative differences between monolingual and bilingual production (Runnqvist *et al.*, 2012), However, variations in the degree of disadvantage have been found to depend on the nature of the words involved (e.g. see Gollan, Montoya & Bonanni, 2005, for evidence of no disadvantage in production of proper names). For that reason, the bilingual disadvantage has been thought to result from less overall use of both L1 and L2 by bilingual speakers: the total amount of time dedicated to each of the language is less than in monolingual speakers, merely because their time has to be split between two or more languages. This has led to the *frequency lag* hypothesis (Ivanova & Costa, 2008; Runnqvist, Gollan, Costa & Ferreira, 2013) or the *weak link* hypothesis (Gollan *et al.*, 2005; Gollan, Montoya, Cera & Sandoval, 2008), a perspective that focuses more on the quantitative differences between bilinguals and monolinguals, in both the production of the L1 and the L2.

In general, evidence points towards a perspective in which monolingual and bilingual lexical production differ mainly from a quantitative point of view, with multiple conditions influencing it (Runnqvist *et al.*, 2012). However, this perspective is focused on lexical production (see Kroll & Gollan, 2014, for a recent review of the findings in bilingual speech production at the lexical level), with much less work attempting to extend these findings and the ensuing debate to production of full sentences. Thus, it could be the case that even if differences in lexical retrieval are a matter of quantity, retrieval of rules in grammatical encoding might differ between languages[21].

A first attempt to analyze syntactic disadvantage in L2 production is the study run by Runnqvist *et al.* (2013). In this study, it was found that, in the production of syntactic structures, there was an influence of the L1 on the L2 even in speakers who use their L2 more often in their daily life. In their study, they had Spanish-English and Mandarin-English bilinguals construct sentences with lexical items that were provided, thus urging speakers to produce active/passive sentences (e.g. "the woman pushes the stroller" and "the stroller is pushed by the woman") and prenominal/postnominal possessive noun phrases ("the woman's stroller is pink" and "the stroller of the woman is pink"). Importantly, frequency of these structures is not the same between languages: passive sentences are less frequent in Mandarin and Spanish than in English; similarly, prenominal possessive NPs are more common in English than postnominal NPs, while prenominal position is the only one that is grammatical in Mandarin, and postnominal the only grammatical form in Spanish. Their results showed that, in general, there was a bilingual disadvantage in whole sentence production, but this disadvantage interacted with frequencies of the structures across languages. In this sense, bilinguals of both languages showed an increased disadvantage in passive sentences compared to the disadvantage with active ones. More importantly, there was a difference between Mandarin-English bilinguals and Spanish-English bilinguals in speech onset latency of possessive NPs. Mandarin-English bilinguals did not show any disadvantage (compared to monolinguals) when producing prenominal possessive NPs, as it is the only available construction in Mandarin. Spanish showed disadvantage in both prenominal and postnominal possessive NPs. In conclusion, these results point to the idea that bilingual speakers rely on their L1 when they face the task of constructing and producing a sentence. Moreover, they point to the idea of disadvantage in L2 production, showing longer latencies and an overall production patterns that differ (at least temporarily) from

[21] This outcome could be predicted by both De Bot's model (since the lexicons, but not the formulators, are shared across languages) and Ullman's model (since lexical retrieval of both languages is performed by the same cognitive mechanisms (i.e. declarative memory), whilst grammatical encoding is undertaken by different mechanisms in L1 and L2 (i.e. by procedural memory in L1 and declarative memory in L2). However, Hartsuiker and colleagues would predict quantitative differences exclusively, since lexical retrieval and grammatical rules are all interconnected in the same language network.

monolingual speakers. Although, more studies would be needed to explore the issue of the differences between the way monolingual and bilingual speakers undertake grammatical encoding, this study seems to point to a similar difference to that found in lexical production.

Qualitative vs. quantitative differences in bilingual sentence processing

The field of language comprehension also provides interesting insights regarding the differences between monolingual and bilingual parsing strategies (Roberts, 2013). Similarly as in lexical production, there is a current debate whether differences between both types of speakers go beyond simple quantitative differences. Clahsen and Felser (2006a, b) argue so. According to the so-called *Shallow Structure Hypothesis* proposed by these authors, although bilingual speakers do not have problems in parsing local dependencies, processing of non-local dependencies that call for complex online processing across clauses seems not to be available for them. This is so even for bilinguals who are highly proficient in their L2.

Actually, several studies have found differences in the ways bilinguals and monolinguals process L1 and L2. Felser, Roberts, Marinis & Gross (2003) analyzed processing of relative clauses with ambiguous attachments, like the one in (1) by German and Greek native speakers with English as L2.

(1) Someone shot the servant of the actress who was on the balcony.

Critically, while native speakers of English commonly show preference for the NP2 ("the actress") as the antecedent of the relative clause (i.e. the actress was on the balcony), thus showing a preference for *Late Closure* (Frazier, 1979; Frazier & Clifton, 1997; Cuetos & Mitchell, 1988), both German and Greek show an attachment preference towards the NP1 ("the servant") as the antecedent (i.e. the servant was on the balcony), with a tendency to assign the relative clause to the noun that is closer to the main clause (*Predicate Proximity*) (see similar results in German: Hemforth, Konieczny & Scheepers, 2000; and Greek: Papadopoulou & Clahsen, 2003).

In their study, Felser *et al.* had offline and online measures of attachment preferences in L2 English. They compared sentences that disambiguated both towards the first NP and the second NP, with genitive constructions (2a) and constructions with preposition “with” (2b).

(2)

(a) The dean liked the secretary of the professors who was / were reading a letter

(b) The dean liked the secretary with the professors who was / were reading a letter

Both German-English and Greek-English bilinguals showed a preference for the NP2 when the construction with “with” was used. However, when the genitive form “of” was used there was no preference for either interpretation. This result was observed in both online and offline tasks, suggesting that bilingual speakers do not transfer their L1 preferences into the L2 in syntactically complex structures. However, more strikingly, Papadopoulou & Clahsen (2003) showed that bilingual speakers did not show any attachment preferences even when monolingual attachment preferences of their L1 and their L2 agreed. Thus, they presented the same type of sentences in Greek to a group of Spanish, German and Russian native speakers who were advanced learners of Greek. Not only do German and Greek favor an NP1 attachment, but also Spanish (e.g. Cuetos & Mitchell, 1988; Carreiras & Clifton, 1993; 1999; Carreiras, Salillas & Barber, 2004) and Russian (e.g. Sekerina, 2003) do so. Despite the congruence between parsing strategies between languages, learners did not transfer their parsing strategies, and showed no attachment preference whatsoever. The results of this and the previous study led authors to suggest that bilingual speakers’ parsing strategies are different in nature from those of monolingual speakers. These strategies converge regardless of typology and the distance with the L1 (see Roberts, 2013 for a discussion).

ERP studies point to a similar outcome, with bilinguals being able to process online semantic anomalies but not syntactic ones. For example, in a study with highly-proficient Japanese learners of German, Hahne and Friederici (2001) found that brain-responses to syntactic and semantic anomalies differ between L2 learners of German and German native speakers: semantic anomalies showed a similar pattern of brain responses (with the appearance of a similar N400 component –although not identical), but the results showed a lack of P600 in L2 learners’ brain response when encountering syntactic anomalies. This evidence, along with spurious appearance of P600 in control sentences by L2 learners, showed a greater difficulty of syntactic integration by advanced learners. Converging evidence comes from an eye-tracking experiment that analyzes the disambiguation patterns for English-French bilinguals and French monolingual speakers (Frenck-Mestre, 1998; in Frenck-Mestre, 2002). In this study, reduced relative clauses[22] were used to analyze whether bilinguals project their native language syntax onto the L2. Thus, participants read sentences like those presented in (1). (1a) is grammatical in both English and French, but (1b) is ungrammatical only in English, the participants' L1:

(1)

a. Le sous-marin **détruit** pendant la guerre a coulé en quelques secondes.

The submarine **destroyed** during the war sank in a few seconds.

b. Le sous-marin **détruit** pendant la guerre un navire de la marine royale.

The submarine **destroyed** during the war a ship from the royal navy.

Results showed that the prepositional phrase "pendant la guerre" blocked the active interpretation only for English-French bilinguals. Bilinguals took longer to read the prepositional phrase region than French native speakers, since that was the disambiguating point for learners (similarly to the English case). Moreover, bilinguals had considerably more processing difficulty (considerably longer reading times) when reaching the object "un navire" in the main clause reading, suggesting that they had previously discarded that interpretation.

This evidence has been seen as an indication of a qualitatively different parsing of bilingual speakers in comparison with monolingual speakers (Clahsen & Felser, 2006a; see Juffs & Rodríguez, 2015, for a review of the evidence supporting the Shallow Structure Hypothesis; but see Dekydtspotter, Schwartz & Sprouse, 2006, for a review of evidence on the contrary).

Use of semantic information in bilingual sentence processing

Although bilinguals have been found to have difficulty with online processing of syntactic complex structures, different studies show that they are able to use semantic and plausibility information in their online parsing. For example, Williams, Möbius and Kim (2001) carried out a study with Chinese, German and Korean learners of

[22] Reduced relative clauses are ambiguous due to the dropping of the relativization maker. Parsers do not know whether they are reading a main clause or a subordinate clause until the main verb appears. The example given in (1a) serves to illustrate this.

English. In this study they manipulated the plausibility in *wh*-constructions, like the ones in (5) and (6).

(5) *Plausible (at V):* Which girl did the man push the bike into late last night?

(6) *Implausible (at V):* Which river did the man push the bike into late last night?

In (5) "the girl" is presented as a plausible object for the man pushing, while in (6) that interpretation seems unlikely, since a river cannot be pushed. A self-paced reading task showed that both monolingual and bilingual speakers had difficulties in (5) upon realizing that the plausible gap for the filler was not correct, leading them to a garden-path situation. Thus, online information of semantic plausibility was used similarly by both native and non-native speakers. However, non-native speakers showed more lasting effects of the garden-path situation, experiencing greater difficulty than monolinguals in recovering from misanalysis.

From a second language acquisition perspective, the role of semantics, as a measure of world knowledge shared between languages, has been considered important for L2 learners: when the knowledge of the code fails, L2 learners may overcome this difficulty by using their world knowledge about thematic roles and animacy combinations. Based on this assumption, the Competition Model (Bates & MacWhinney, 1982; MacWhinney, 1997; 2001) proposes that language comprehension is guided by different cues (syntactic: word order, case assignment; morphological: verbal and/or nominal morphology; semantic: animacy of the elements involved, etc.) that compete with each other. Not all languages assign the same strength to these cues, resulting in cross-linguistic differences between them[23]. Thus, L2 learning involves acquiring the proper cues of the target language, along with its balance in the competition system.

Studies from the Competition Model approach (see MacWhinney, 1997 for references of studies carried out with different languages) have tried to analyze whether L2 learners use in their interlanguage[24] the proper L2 cues or, to the contrary,

---

[23] The strength of particular cues depends on the probability that a specific cue will lead (or has led in previous experiences) to the correct interpretation of the sentence (MacWhinney, 2001).

[24] The interlanguage (hereafter, IL) is the linguistic system that characterizes L2 speakers. It contains both elements of the L1 and of the L2, but it is idiosyncratic, and contains elements that do not belong to either language (Selinker, 1972; See also Gass & Selinker, 2008).

they maintain their L1 system and apply it incorrectly to L2. Results suggest that in the first stages of IL development, L1 cues are maintained and used wrongly when parsing L2 sentences. Thus, for example, Sasaki (1994) found that L2 Japanese learners (L1 English) were using English cues when interpreting sentences, as they based their interpretations on word order[25]. However, as learners attained more proficiency in Japanese, they shifted to L2 cues (essentially, case particles). Competition Model studies point to changes in the use that learners make of language cues, turning gradually into the proper L2 cue strength. However, in this restructuring process, the learner will pass through an intermediate level of cue adaptation, in which they will tend to rely on a universal strategy based on meaning and on the semantic values of the utterance, animacy being a strong guide for interpretation (Gass & Selinker, 2008)[26]. In conclusion, the Competition Model will predict that when processing sentences in the L2, there will be an early stage in which bilinguals rely heavily on their L1. This stage will give way to a universal reliance on conceptual relations between elements, before bilinguals are able to settle and use native-like cues accordingly. L1 cues will be maintained to different degrees depending on languages: English native speakers, for example, easily drop word order as a cue, relying from early stages of development on semantic relations between elements. However, speakers from languages with more free word orders (like Italian or Japanese) will not adopt word order cues early in the process, maintaining their own native language cues for a longer period (Gass & Selinker, 2008).

Several studies from this perspective have tried to analyze the changes in linguistic weights from L1 to L2, as the learner advances in her IL. These studies (Gass, 1987; Harrington, 1987, both in Kanno, 2007; Sasaki, 1994) show the predominance of a strategy based on semantics, in particular the use of animacy cues, in L2 sentence processing and IL development. These semantic cues often precede grammatical cues (e.g. word order). Kanno (2007), in a study of RC comprehension in Japanese as L2, observed that reversible RCs (i.e. RCs wherein both antecedent and

---

[25] A cue that is important in English, due to its rigid word order, but that is not informative in the case of Japanese, a language that allows *scrambling* (Tsujimura, 2007).

[26] However, Gass and Selinker (2008) noted in their review that this picture is not so clear, as the strength of the reliance on animacy and the exact moment at which learners abandon their native language cues vary according to both L1 and L2 (see Gass & Selinker, 2008, for a more detailed review).

subordinate HN are animate and the verb is reversible, for example *This is the dog that chased the cat*) had a correct response pattern close to chance level, while non-reversible sentences increased significantly the amount of correct responses. The author concludes that the scarce experience of her participants with Japanese RC could have led them to be unable to use Japanese cues, like case marking. When L2 cues are unknown, cues provided by semantics are essential in order to perform the task in a correct way. Thus, learners will make use of animacy from early moments of L2 development, giving it priority over other cues under certain circumstances.

Whether L2 learners can make use of animacy cues to improve their comprehension has been of special interest in the field of RC comprehension. As was presented in Chapter 1, Japanese RCs show special features, in that some RCs in Japanese do not possess a gap-filler dependency (Shirai & Ozeki, 2007) and are based completely on semantic relations between the sentence constituents (Tsujimura, 2007). If this were the case, the cues provided by animacy should be essential to interpret relative clauses in this language, thus accounting for the early effects of animacy on RC comprehension. However, similarly to the study by Kanno (2007), other authors, with different languages, have tried to analyze if, through the manipulation of animacy combinations, the differences between subject and object RCs will disappear or, at least, be reduced.

Jackson and Roberts (2010) also show evidence of the use of animacy information by German-Dutch bilingual speakers during the processing of relative clauses with an animate or an inanimate head noun, like the ones in (7a) and (7b) and (8a) and (8b):

(7)

(a) Vor de kinderen is de clown, die de taarten *heeft* gegooid, het hoogtepunt van de voorstelling. (Subject RC, animate subject)

For the children is the clown$_{SG}$, that the pies$_{PL}$ has$_{SG}$ thrown, the highlight of the performance.

“For the children the clown, that threw the pies, was the highlight of the performance”

(b) Vor de kinderen is de clown, die de taarten *hebben* geraakt, het hoogtepunt van de voorstelling. (Object RC, inanimate subject)

For the children is the clown$_{SG}$, that the pies$_{PL}$ have$_{PL}$ hit, the highlight of the performance.

"For the children the clown, that the pies hit, was the highlight of the performance"

(8)

(a) Vor de kinderen zijn de taartem, die de clown *heeft* gegooid, het hoogtepunt van de voorstelling. (Object RC, animate subject)

For the children are the pies$_{PL}$, that the clown$_{SG}$ has$_{SG}$ thrown, the highlight of the performance.

"For the children the pies, that the clown has thrown, was the highlight of the performance"

(b) Vor de kinderen zijn de taartem, die de clown *hebben* geraakt, het hoogtepunt van de voorstelling. (Subject RC, inanimate subject)

For the children are the pies$_{PL}$, that the clown$_{SG}$ have$_{PL}$ hit, the highlight of the performance.

"For the children the pies, that the clown have hit, was the highlight of the performance"

Their results showed larger reading times in both groups (native and non-native) for sentences with the animate head noun in which the interpretation pushed towards an object RC (i.e. (7a) larger reading time than (8a)). This suggest that participants were processing online the animacy cues and creating a structure according to that information. On the other hand, and similarly to other studies run only with monolingual speakers (e.g. Mak *et al*., 2006 –see chapter 2), both native and non-native participants did not show any parsing preference when the head noun (subject) was inanimate ((7b) and (8b)).

These results show that, under certain circumstances where the proper semantic information is given, bilingual speakers are able to process exactly like monolingual speakers, arguing against a qualitative difference between both types of parsers, and showing only quantitative differences. Semantic information is used by bilingual speakers in order to ease processing costs brought about by parsing in an L2, costs that can underlie the differences found between native and non-native processing.

Other factors affecting bilingual sentence processing: Working memory and language proficiency

Processing in an L2 comes with spurious activation, and increased working memory and attention costs that might result from the lack of automatization of many of its features (Segalowitz, 2003). Working memory capacity has been shown to affect parsing strategies in monolingual speakers, modifying the way they parse complex sentences (e.g. Just & Carpenter, 1992; Swets, Desmet, Hambrick & Ferreira, 2007). This extended idea has led to some authors to explore the relation between working memory capacity and L2 processing. Dussias & Piñar (2010) showed that Chinese-English bilinguals were able to use plausibility during processing of *wh*-movement sentences. Both English native speakers and Chinese L2 speakers of English were misled in the first place by the dispreferred parsing of the *wh*-sentences (object extraction v. subject extraction), followed by a reanalysis. However, only bilinguals with high working memory capacity resembled native speakers in their capacity to recover from misanalysis. Moreover, McDonald (2006) found that monolingual speakers with low working memory capacity performed like late bilinguals in their L2 when they had to carry out grammatical judgment analysis under cognitive taxing situations.

However, there are also studies with monolingual speakers showing no correlation of performance in parsing with verbal working memory (e.g. Caplan & Waters, 1999) or L2 memory span (e.g Juffs, 2005; Felser & Roberts, 2007). Working memory is a wide and complex system with different subsystems. Thus, the way WM is measured (Szmalec, Brysbaert & Duyck, 2012), along with the procedure used to analyze its scores (Juffs & Rodriguez, 2015) seem to be crucial in determining the effects that working memory capacity has on language processing. Nevertheless, it seems that, despite complex evidence, spurious activation during sentence processing taxes incremental parsing in bilingual speakers, by delaying it or even blocking processing entirely. Consequently, bilinguals can perform as native speakers if enough working memory resources are available, either by means of a previous higher WM capacity or of gradual automatization of processes that allows freeing working memory resources and devoting them to language parsing activities.

Moreover, experiments by Foucart and Frenck-Mestre (2012) show that these processing differences are not permanent, but rather dependent on L2 proficiency. The authors recorded EEG and eye-tracking data in the processing of noun-gender agreement violations of English-French late bilinguals (with high-proficiency) and French native speakers. The study measured the EEG pattern of three noun-adjective combinations: (1) noun – adjective order, the most prototypical in French; (2) adjective – noun order, less prototypical, shared by both English and French but with different properties, due to required grammatical agreement in French; and (3) adjectives included in a predicate, separated from the noun by the main verb of the sentence: noun – verb – adjective. The results of the EEG study showed that the first condition caused the same activation pattern for both monolingual and bilingual speakers: the appearance of a P600 in gender-agreement violations between the noun and the adjective. Experiments 2 and 3, with less prototypical noun – adjective combinations, however, did not display a P600 for learners, in contrast to the results found with French native speakers: pre-nominal adjectives (experiment 2) caused an N400 response, while copular sentences (experiment 3) did not elicit any significant response. This last condition (i.e. noun and adjectives being separated by syntactic boundaries) was also evaluated with an eye-tracking task, in a fourth experiment. Results showed that, despite the differences in EEG patterns, fixation patterns between both groups were similar. The authors concluded that late learners can process syntactic structures like native speakers, but processing depends on proficiency and on the ease/prototypicality of the structures: structures that are unique in the L2 and cannot be confused with L1 structures are more easily acquired fully[27]. Similarly, Dussias and Sagarra (2007) showed further evidence of the role of proficiency with relative clauses that presented an attachment ambiguity. They carried out an experiment with Spanish-English bilinguals (both low-proficiency and high-proficiency) residents in the U.S., using eye-tracking methodology. Participants had to read relative clauses with double antecedent in their L1 (i.e. ambiguous relative

[27] Regarding the differences between experiment 3 and 4 (EEG and eye-tracking data with the same materials), the authors conclude that they are due to differences in the experimental paradigms: EEG experiments call for word-by-word presentation of the sentences, which is more taxing for working memory. In contrast, eye-tracking allows for a full-sentence presentation, with possibility of backward movements and peripheral vision. Greater working memory load has been suggested to hinder L2 learners performance (McDonald, 2006)

clauses), like examples (3) and (4):

(3) An armed robber shot the sister of the actor who was on the balcony.

(4) Un ladrón armado le disparó a la hermana del actor que estaba en el balcón.

Sentences were disambiguated (by means of morphological and conceptual gender of the constituents) either toward attachment to the first noun ("high-attachment", shown to be preferred by Spanish monolingual speakers) and or to the second noun ("low-attachment", preferred by English monolingual speakers). The pattern of gazes and eye-fixations on L1 relative clauses were different depending on L2 proficiency: low proficient bilinguals showed a high-attachment preference, similar to Spanish monolingual speakers, whilst high proficient bilinguals were affected by their L2 when processing relative clauses in their L1: they displayed a low-attachment preference pattern, thus being affected by their L2 while processing L1, in participants living in their L2 environment and making use of L2 frequently. These results suggest that not only bilinguals can process syntactic features like native speakers when enough exposure to the L2 has been reached, but also that they transfer these parsing preferences to their own L1.

Producing sentences: relation between L1 and L2

The evidence presented so far, points to a relation between L1 and L2 that can go in both directions. Although there are differences in sentence parsing between monolingual and bilingual speakers[28], these differences seem to be mediated by language proficiency and cognitive resources.

Research on full sentence production by bilingual speakers is much more reduced, and much of it comes from the Second Language Acquisition field. Despite the risks that entail generalizing language processing findings to the field of language production (Costa & Santesteban, 2004b)[29], studies on language production show

[28] As Kroll & Gollan (2014) point out, the difference should be drawn between monolingual speakers and bilingual speakers, rather than between L1 or native speakers and L2 speakers, since the effects of bilingualism seem to affect both bilinguals' languages, and go beyond the realm of language.

[29] Litcofsky, Tanner & van Hell (2016) show that even when lexical processing and production are highly related in the L1, this relation is not so straightforward in the L2, due to the more extensive variation in L2 in comparison with L1 in terms of neural activity. These authors found that production was more related to L2 proficiency, whilst they failed to find correlation with working memory for production, a factor that was focal in predicting L2

similar outcomes regarding the relation between L1 and L2, and the use of semantic information by bilingual speakers.

Different studies that have made use of structural priming methodologies or sentence completion tasks have consistently reported a strong influence of the L1 over the L2 and viceversa (e.g. structural priming: Hartsuiker *et al.*, 2004; Meijer & Fox Tree, 2003; Salamoura & Williams, 2007; Shin & Christianson, 2009; sentence completion: Hoshino & Kroll, 2010; Hatzidaki, Branigan & Pickering, 2011). These studies, as will be reviewed in more detail below, show that there can also be an influence of L2 on L1. This influence, however, is highly dependent on L2 proficiency (e.g. Schoonbaert, Hartsuiker & Pickering, 2007), and on the kind of structure under analysis (i.e. relation between L1 and L2 in that structure in particular) (e.g. Bernolet, Hartsuiker & Pickering, 2007; Loebell & Bock, 2003). Exploring this directionality of the relation between languages, Hatzidaki *et al.* (2011) reported a study with Greek-English and English-Greek bilinguals in which they analyzed grammatical number activation in a sentence completion task. Participants were presented with an NP and asked to continue the sentence (the NP was at the beginning of the sentence). Critically, the grammatical number of some of these NPs differed between Greek and English, and hence the number of the verb to be used. For example (from Hatzikadi *et al.*, 2011), "the money" is a singular noun in English, but holds a plural number in Greek ("divergent condition"). Results showed that the non-target language was also activated throughout the task, as there were more errors in divergent conditions (when the number in both languages differs) than in convergent conditions (when the number in both languages is the same), where there were no errors in general. This effect was stronger from the stronger L1 to the weaker L2, but some effect of L2 grammar over L1 grammar was also found, especially when the context favored activation of both languages ("bilingual mode"; De Groot, 2011) (see Pavlenko, 2000, for a review of studies that show effects of L2 over L1 in late bilinguals). More recently, however, Hartsuiker *et al.* (2016) showed that, once "enough" proficiency is reached, priming effects from L1 to L2 do not differ from effects from L2 to L1 (and also from effects between different L2s and within languages).

---

comprehension (but not so much proficiency level). Note, nonetheless, that Linck, Osthus, Koeth & Bunting (2014) found a correlation between working memory and L2 production (and comprehension) in their meta-analysis study.

Taken together, these results suggest that not only is the grammar of both languages closely related, but also that the grammar of the other language cannot be deactivated when speaking, reaching even a point at which both languages have equivalent repercussions on each other. However, this relation does not hold invariantly and depends not only on language proficiency, as Hartsuiker *et al.* (2016) note, but also on language environment. One example of this was given above, with the Hatzidaki *et al*.'s study, where the experimental environment (whether the task was performed in two languages or only in one) affected the degree of activation of the non-target language, not only the L1 but also the L2.

Similarly, Hoshino, Dussias and Kroll (2010) showed further evidence of the relation between L1 and L2 syntax regarding L2 proficiency with Spanish-English and English-Spanish bilinguals at two proficiency levels. They analyzed subject-verb agreement with a sentence completion task, in which participants had to continue the beginning of sentences like the following (from Hoshino *et al.*, 2010):

(4) The author of the novels (Single-Referent, Number Mismatch): Grammatical number.

(5) The drawing on the posters (Distributive-Referent, Number Mismatch): Conceptual number.

Interestingly, results showed that low-proficiency bilinguals were sensitive only to grammatical number, but could not make use of conceptual number. Only high-proficiency learners were affected by both types of number. Note that the study by Hahne and Friederici (2001) with ERP reviewed before pointed to a greater difficulty in using syntactic information by L2 speakers, but to a reliable use of semantic information, thus showing differences due to linguistic modality, which are worth taking into consideration. However, Hahne and Friederici's study was a comprehension study, while Hoshino *et al*. tried to analyze production patterns. Both comprehension and production processes have different departing points and different objectives (i.e. transforming conceptual information to a linguistic message in language production, and decoding a linguistic message in order to be able to access the conceptual information encrypted on it in the case of language comprehension). Studies on lexical access by bilingual speakers have shown that these differences have consequences in the type of information that affects each of the processes, and the strength with which this information is involved (e.g. semantic or syntactic information) (see De Groot, 2011, for a review of bilingual lexical processing – comprehension and production– studies).

An incremental bilingual production system

There are a few central questions in language production that are worth considering in the analysis and comparison among different studies. Bilingual language parsing takes places incrementally, although extra cognitive load and high complexity of sentence structure can delay parsing until the end of the sentence (Roberts, 2013). Similarly, sentence planning by bilingual speakers takes place in an incremental fashion, and is constrained by time limits caused by real time speech. In this planning process, similarly to monolingual speakers, more accessible items will tend to be processed and produced earlier than less accessible ones (see the review of monolingual production in the previous chapter of this dissertation). The bilingual's linguistic system, however, has to deal with spurious activation, e.g. phonological and lexical activation (Broersma, 2005), semantic activation (Talamas, Kroll & Dufour, 1999), and greater cognitive load due to the effects of L1 and/or non-automatized processes (Ardila, 2003). This results in a somewhat vague picture, in which many different aspects can affect the filtering of more accessible information. The effects of accessibility on bilingual sentence production will be reviewed in the following sections, roughly using the same distinctions of the previous chapter, namely, structural accessibility, conceptual accessibility, planning scope and timing of production planning.

Structural accessibility

As it was described in the previous chapter, more readily accessible constructions will tend to be more easily activated and, thus, produced earlier, or preferentially, than less accessible ones. Specific constructions can be activated due to frequency in language, canonical prototypicality, or merely owing to previous use or exposure to that particular syntactic structure (Hartsuiker & Pickering, 2008). Structural priming shows, as it was shown in chapter 2, that only structurally, but not superficially, similar constructions will be favored by the effects of accessibility. This occurs as a result of remaining activation of previously used constructions: similar syntactic constructions are interconnected in the speaker's mind. Thus, cross-linguistic structural priming studies provide a way to analyze whether there are cross-linguistic connections and the extent of those connections between L1 and L2. For example, if both language structures are connected, when a Spanish-English bilingual produces a passive sentence in Spanish, the probabilities of producing subsequently a passive sentence in English should be greater than when producing an active sentence. Hartsuiker *et al.*'s (2004) study shows that this is indeed the case. Spanish-English bilinguals were asked to describe a picture in their L2 after a priming trial: they tended to produce English passives more often after using a Spanish passive than active or intransitive sentences. Similar evidence of priming between Spanish and English has been found with dative sentences (Meijer & Fox Tree, 2003), in an experiment in which Spanish L1 – English L2 participants had to recall sentences after repeating primes. Results showed that bilinguals were more likely to misremember dative sentences with double-object structure than datives with a prepositional object after Spanish sentences containing this structure.

Cross-linguistic priming has been observed across different language combinations: priming in dative sentences from L1 to L2 by German-English bilinguals (Loebell & Bock, 2003)[30], and Greek-English bilinguals (Salamoura & Williams, 2007). Additionally, Schoonbaert *et al.* (2007) found evidence of bidirectional cross-linguistic priming (i.e. from both L1 to L2 and L2 to L1) with Dutch-English bilinguals. Interestingly, this also has been observed in relative clause attachment (high and low attachment for ambiguous sentences[31]), with Dutch-English bilinguals (Desmet and Declercq, 2006) and with Dutch-English-French bilinguals, between all their languages (Hartsuiker *et al.*, 2016).

Regarding the influence of word order in cross-linguistic priming, the evidence is mixed. Bernolet *et al.* (2007) carried out an experiment using a confederate-scripting task[32] with German-Dutch bilinguals and Dutch–English

[30] These authors, however, did not find comparable results when analyzing transitive sentences.

[31] See chapter 1 for a brief review and examples of this type of relative clauses.

[32] The confederate-scripting task consists of a picture-description game: two persons (the confederate and the participant) describe pictures to each other, either in the same language (intralinguistic condition) or in different languages (interlinguistic condition). The type of syntactic construction that the confederate uses at each moment is controlled. Subsequently,

bilinguals with relative clauses. German and Dutch RCs share the same word order. However, the word order of RCs for those two languages differ from that of English, as can be seen in (6) (from Bernolet *et al.*, 2007; English gloss added):

(6)

a) the shark that is red (RC-structure, English)

b) de haai die rood is (RC-structure, Dutch)

(the shark that red is)

c) der Hai der rot ist (RC-structure, German)

(the shark that red is)

These authors observed a priming effect in the German-Dutch group (i.e., bilinguals whose languages share the same word order), but not in the Dutch-English group (i.e., bilinguals whose languages do not share relative clause word order). They attributed this difference to the effect of word order. According to these authors, same word order is a requisite for sharing nodes between languages, with grammatical encoding undertaken in a single stage containing both grammatical function assignment and word order assignment (see chapter 2 for more details about grammatical encoding processes). However, Shin and Christianson (2009) found cross-linguistic priming with Korean-English bilinguals in dative sentences (by using a sentence recall task), like the ones presented in (7) below, in which word order between languages differs (from Shin & Christianson, 2009; explanations and English translation added):

(7)

a) Mary-ka John-eykey chayk-ul cwu-ess-ta

Mary-NOM John-to book-ACC gave-PAST-DECL (Postpositional dative construction): Canonical form.[33]

---

the degree to which the participant tends to repeat same syntactic construction than the confederate used is measured (Branigan, Pickering & Cleland, 2000).

[33] There is also a scrambled version of the postpositional dative construction, in which the accusative “book” precedes the dative “John”. However, the authors mention that it was not accepted by a majority of native speakers in a pilot study, which led them to discard this structure in their study.

"Mary gave a book to John".

b) Mary-ka John-ul chayk-ul cwu-ess-ta

Mary-NOM John-ACC book-ACC gave-PAST-DECL (Double-object dative construction): less used structure, restricted to few contexts.

"Mary gave John a book".

Korean, similarly to Japanese, is a verb final language, which is consequently formed by postpositional structures, in contrast to English prepositional structures. However, as the authors comment, Korean and English double-object constructions (7b) (clearly an unpreferred structure in Korean) share the same argument order and functional-level structure, while pre/post-positional dative constructions (7a) share only the same argument order. Priming effects were limited to double-object sentences.

A recent study by Rodrigo, Tamura, Kubo, Tanaka & Koizumi (2016) with Spanish-Kaqchikel bilingual speakers shows similar structural priming effects regardless of word order differences. In this study, we controlled the effects of grammatical assignment and word order in the relation between languages. Kaqchikel's canonical word order is VOS, although SVO is also widely used. Similarly, Spanish preferred word order is SVO, but its flexibility allows VOS word order as well. In a confederate-script task, participants listened to sentences with SVO or VOS order, in active or passive voice in Spanish (their L1). Results showed there was an increase in passive sentences in Kaqchikel (L2) after passive primes, regardless of word order (i.e. word order differences did not block structural priming effects), but there was no priming effect of word order alone (VOS primes did not cause an increased number of VOS responses in Kaqchikel). However, priming within Kaqchikel showed effects of both voice and word order, as it was shown in chapter 2. These effects suggest that same word order is not a mandatory requisite for languages to establish connections between different grammatical functions, pointing more towards a two stage vision of grammatical encoding. The two languages in a bilingual's mind are connected at the level of grammatical function assignment, but this connection is no longer present (or at least it was not found in our experiments) at the constituent assembly level.

In conclusion, it seems that, despite the fact that intralinguistic structural priming exhibits consistent results across languages and tasks, cross-linguistic priming effects are more varied, and seemingly weaker, showing a sensitivity to a number of variables that can minimize them or even make them disappear: type of task, L2 proficiency, etc. (Shin & Christianson, 2009; Shin, 2010). However, what can be concluded from these studies is that there is a relation between the L1 and L2 formulators, although the connection points[34], the strength of the connections[35], and the extent to which other variables (like the ones listed above) can affect this relation are still open to discussion.

Conceptual accessibility

Some studies on bilingualism have tried to examine the use of conceptual information that learners show when they listen to/read an L2 sentence or speech. However, this is not the case for production studies, since there are virtually no attempts, as far as we know, to analyze the mapping processes between conceptual information and the formulator with bilingual speakers. The Competition Model (Bates & MacWhinney, 1982), explained in the previous section, is not only focused on how sentence processing takes place, but also applies the same idea of semantic cues as a kind of universal features that facilitate speech production when L2 cues are still unknown. This leads to the prediction that bilinguals will heavily rely on animacy in early to intermediate stages of the IL development.

Ozeki and Shirai (2007) analyzed the role of animacy in the oral production of subject- vs object-RCs by L2 learners. In their study they tried to test the prediciton derived from Kanno (2007) (see above) that Japanese RC learning relies heavily on animacy. They carried out a study on the use of RCs by Japanese L2 learners with different L1s (English, Korean and Mandarin). Through the analysis of oral

[34] Whether languages are connected at the grammatical function level in a two stages model (Shin & Christianson, 2009; Rodrigo *et al.*, 2016) or the relation takes place in a single stage with nodes that contains information of both grammatical function and word order (Bernolet *et al.*, 2007).

[35] As it has been reviewed before, different theories make different predictions about the strength of cross-linguistic priming, in comparison to same-language priming. Hartsuiker *et al.*'s theory is the only one that assumes that the strength of both between and within language priming should be the same, and independent of proficiency. In 2007 they reviewed several works that found similar strengths in both L1 and L2 structural priming effects. However, the picture is still not clear; many studies have failed to show priming from L1 to L2, and the relation seems to be highly dependent on proficiency (e.g. Loebell & Bock, 2003 for transitive sentences).

interviews, they found that L2 learners, even at low proficiency levels, were able to produce object RCs, even to a greater extent than subject RCs. As learners attained proficiency in Japanese, the amount of subject RCs increased, thus resembling the distribution shown by Japanese native speakers (who showed preferences for subject RCs –see Chapter 1). When they analyzed the role of animacy in learners' spoken responses, they found that learners displayed a strong relation between subject RCs and animate antecedents, and object RCs and inanimate antecedents. This pattern was attributed to the fact that learners might have been guided by the animacy of the antecedents rather than by grammatical relations. Nonetheless, Japanese native speakers made associations between the animacy of the antecedent and the type of relative as well, but this correspondence was different to that found in L2 learners: RCs with animate antecedents were produced by native speakers mainly with the animate item taking the subject function. However, the opposite relation did not hold: not all the produced subject RCs were constructed with animate antecedents. In fact, subject RCs with inanimate antecedents were produced in approximately half of the cases. On the contrary, the association pattern shown by learners was different: they showed a biunivocal association between animate antecedents and the subject function of the relative clause. In other words, almost all the sentences constructed with animate antecedents had the animate item as the subject, and at the same time most subject-RCs modified an animate antecedent. As for object RCs, however, this bidirectional relation is found for both natives and learners: both cases, most RCs with inanimate antecedents were assigned the object function, and the majority of object RCs were produced with inanimate antecedents.

This pattern of associations between animacy and RC type led the authors to carry out a second study, in which participants had to combine two given sentences into an RC in a written test, in which the animacy of the antecedent was controlled for. Participants were intermediate and advanced Cantonese speakers learning Japanese as L2. Although they did not find a general effect of animacy in a global analysis of responses, an analysis of the errors made by participants showed that lowest level learners were correctly producing subject RCs with animate antecedents only, and that most errors consisted of changing object (and oblique) RCs to subject RCs when the antecedent was animate. Since this error only showed up in sentences with animate antecedents, the authors concluded that there is a strong relation between animate antecedents and subject RCs for L2 learners. In conclusion, this study showed that animacy plays an essential role in the production of RCs by learners of Japanese, who tend to overstate the association between animate antecedents and subject RCs, as though it were established by rule.

Nevertheless, from these results it cannot be ascertained whether this is a special characteristic that defines Japanese RCs or not. To be sure, Japanese might have a system that favors a relation between elements based on semantics exclusively, and hence displaying stronger or clearer effects of animacy on L2 learning. However, this is not the only language in which this dependence on semantic information has been found, as it was shown extensively in the previous section and in Chapter 2. (e.g. L2 comprehension: Dussias & Cramer Scaltz, 2008; Felser, *et al.*, 2003; Papadopoulou & Clahsen, 2003; Frenck-Mestre, 2005; Clahsen & Felser, 2006b; L1 comprehension Townsend & Bever, 2001; Mak, Vonk & Schriefers, 2002; 2006; Traxler, *et al.*, 2005; MacDonald, Pearlmutter & Seidenberg, 1994). Japanese shows stronger reliance on animacy cues than western languages (Rosen, 1999)[36], which might lead bilinguals to follow conceptual relations as the guide for sentence planning more closely and/or at an earlier stage of development than bilinguals of other languages, thus providing a suitable opportunity to explore accessibility effects on bilingual sentence production.

For that reason, Rodrigo (2013) presented a study on RC acquisition by Spanish native speakers who were learning Japanese, with three levels of proficiency (beginners, intermediate and advanced). In this study, I analyzed the role that animacy plays in RC production and how the relation between L1 and L2 evolves with proficiency. Results showed that animacy cues were extensively used at intermediate levels, but not so at low levels. Word order and basic syntactic features were acquired early, despite the stark contrast in word order between the two languages, and the only errors that remained was a tendency to assign subject case particles to animate elements regardless of the scene and the verbal voice used. These results show a central role of animacy in bilingual sentence production.

[36] Rosen divides languages according to the linguistic representation they do of linguistic events. Japanese, along with Icelandic or some ergative languages, belongs to the group of I-languages (initiation languages, in contrast with D(elimitation)-languages, like English). In this type of languages there is greater sensitivity to the properties of the subject, with a strong constraint in the animacy features it can hold.

In conclusion, these two sections have shown that (1) the languages of a bilingual speaker are related to each other and affect bilingual language production. This relation is found at the level of grammatical function assignment, within grammatical encoding processes (and of course, in lexical access). (2) Conceptual information, in general, and animacy information, in particular, is used as a guide for sentence production, helping bilinguals to create mappings that conform to prototypical animacy combinations. This was denoted by the type of errors that bilingual speakers make and by their overreliance on animacy as a cue to decide the structures to be produced. However, these studies do not show clearly whether these factors operate incrementally or are the result of reanalysis, inducing a more serial planning process for bilingual speakers in comparison to monolingual speakers. In the next, and final subsection, we will address this question, by analyzing the only study we know so far that makes use of eye-tracking methodology to explore online sentence production by bilingual speakers.

Online bilingual sentence production: Contributions from eye-tracking methodology.

Increased proficiency in an L2 is usually associated with an increased fluency in normal speech (De Jong, Steinel, Florijn, Schoonen & Hulstijn, 2012). Similarly to monolingual speakers, bilinguals are faced with the task of producing speech in a pace that allows for conversation to unfold without major breaks. This process comes by way of increased experience with the target language, which allows for faster processing (Konopka & Forest, 2016).

Despite the problems of automatization of working memory load that bilingual speakers must cope with, speakers with more than one language are able to plan their utterance in an atomic way, by planning only one lexical item at a time under certain circumstances. This was the results reported by Brown-Schmidt and Konopka (2008). This study comparing Spanish and English speech planning order and timing was presented in the previous chapter. The authors reported the planning differences as cross-linguistic differences between the two languages. However, and more important for the purposes of this chapter, their participants were not monolinguals of both languages but the same participants, who were bilinguals, undertook the same task. In their task, as previously reported, participants had to identify an object in English or in Spanish by using a size adjective that allowed to contrast it with another element (e.g. “the small butterfly”, when they had to contrast it with a bigger butterfly). In Spanish, the adjective takes a post-nominal position, in contrast with the prenominal position in English. Their results showed that gazes to the contrasting element (i.e. the big butterfly) were delayed in Spanish in comparison with English, thus showing a planning process that took place incrementally word by word. These results, then, show that bilingual speakers are not only able to plan their speech incrementally, but also that this process in flexible: the very same bilinguals changed the order of their gazes accordingly with the grammar of the language on hand.

Note, however, that the participants in the Brown-Schmidt & Konopka study were early bilinguals, and the task on hand did not call for complex syntactic structures, but only NPs with an adjective. What happens when late bilinguals are faced with the task of planning a full sentence in their L2? Do they plan their speech according to a linear incremental planning (as the results from Brown-Schmidt & Konopka suggest), or they rely more heavily on hierarchical incremental planning?

Konopka and Forest (2016) carried out a study in order to explore this question. They had Dutch-English late bilinguals with an intermediate-advanced level of English describe pictures with transitive actions in their L2. Their results showed that the onset of gazes to the first mentioned element was delayed in comparison with gazes when the sentence was delivered in their L1 Dutch. In order to clarify the underlying reasons behind this delay, the authors ran two more experiments. In their experiment 2, they presented participants with a preview of the first word (the subject), while in the experiment 3 participants saw a preview of the verb. Thus, if the reason of the delay was merely a difficulty in accessing the name of the first element, differences between L1 and L2 should disappear after a preview of this name. On the other hand, if the delay in the onset of encoding the first name is caused by a difficulty in encoding the action that is taking place, a preview of the verb should help eliminate this difference.

Their results showed that the preview of the first word did not change the timing of focusing the item corresponding to the first mentioned NP. However, upon preview of the verb, the delay found in onset of gazes to the first NP disappeared. Their results showed that L2 speakers take more time in hierarchical planning than L1 speakers, but this effect is mediated by familiarity with the target verb. Accumulated experience can shape sentence planning by creating a more or less extensive initial planning.

In the previous chapter we presented evidence that showed how speech planning scope depends on available cognitive resources and the ease to access the lemmas of the nouns and verbs of sentences. It is important to remember that bilingual speakers' processing resembles that of low WM capacity monolinguals, and that it can lead bilinguals to plan their L2 like L1 speakers do under high cognitive strain, that is, by reducing their planning scope and making them resort to a word-by-word planning. The Konopka & Forest study seems to suggest this is not the case. However, their participants were producing transitive sentences, a type of utterance that is highly familiar to advanced bilinguals and shares the same word order in English and Dutch. Moreover, the same problem we mentioned in the previous chapter remains here: the first constituent in the sentence happens to be the subject, the most dominant element. It could be the case that a preview of the verb facilitated encoding of the message or the gist of the scene (i.e. what is happening in the scene), but not so much the construction of a structural scaffold that contains the relation between elements. In order to explore these questions, in this dissertation we will present a study with late Spanish-Japanese bilingual speakers, who possess an advanced level of Japanese. We will explore the timing of RC planning in L2 Japanese, as this kind of structure is widely used and at the same time syntactically complex, with significant contrasting features across both languages.

Part 2.

Empirical Studies

Chapter 4. Studies 1 and 2: Relative Clause production in Spanish and in Japanese by monolinguals speakers

The evidence presented along the previous chapters renders a picture of language production as a flexible process in which both lexical and structural factors play an important role. Conceptual accessibility, as well as structural accessibility, affects the product of speech planning in a complex relation between the two (Prat-Sala, 2000; Gennari *et al*., 2012; Tanaka *et al*., 2011; Gennari & MacDonald, 2009; MacDonald, 2013a), in ways that seem to differ across languages, suggesting not only that the selection of the starting point of the utterance is important (Bock *et al.*, 2004), but also the way in which these two crucial pieces of information are handled along the whole process (Bock & Ferreira, 2014). However, it is not clear how speakers coordinate the use of relational (structural configuration of constituents) and non-relational information (access to lemmas) along the whole planning process, and how this changes across different grammatical systems, which posit different constraints on the planning process (Norcliffe & Konopka, 2015).

With this question in mind, we first present two studies run with monolingual speakers of two typologically distant languages: Spanish and Japanese. In these studies we aimed to explore:

1. How different planning stages interact during sentence production. Namely, we explore the use of structural and lexical information along the whole planning process, from conceptual encoding to articulation, by analyzing the strategies used at different points of the process:
   a. We analyze the information that is used first when undertaking linguistic encoding processes (both the left and the right sides –lexical encoding and grammatical encoding– of the Ferreira & Slevc (2007) model presented in the previous chapter), starting from the moment conceptual encoding is carried out. In other words, we will try to establish how the starting point is decided and the role that conceptual accessibility plays on it.
   b. As stated above, both structural and lexical information are essential and closely related during speech planning. However, one of the main questions that remains to be fully understood is which information is prioritized in

speech planning when undertaking the planning of a complex structure. Therefore, we aim to explore whether structural or lexical information is prioritized before speakers start constituent assembly in the same order as the elements are going to be uttered. This point is closely related to the previous one. We expect that the relation of the answers to both questions will allow us to understand how speakers decide to start from the point they do and whether this is a constant process, or depends on the accessibility of concepts or structures.

c. Finally, although speech planning is understood as an interplay of relational and non-relational processes (Konopka & Meyer, 2014), it is still not clear how both types of information are related along the whole planning process. In order to explore this point closer, we aim to analyze whether and how structural information is related to lexical retrieval once constituent assembly process (name-related gazes in the order of mention) has begun.

2. We explore these questions with the aim, as well, of comparing whether and how cross-linguistic variation in grammatical features and word order affects the interplay of stages in sentence formulation. Main differences between Spanish and Japanese were summarized in chapter 1. These differences (mainly in terms of word order and the use of animacy cues) will allow us to analyze the idiosyncratic processes that characterize each language, along with the parts of the process that are common to both.

In order to address these questions, we compare the production of complex structures (relative clauses –hereafter RCs) in Spanish and Japanese (both RCs with the agent and with the patient as the head noun -hereafter HN) by means of the eye-tracking methodology in a visual world paradigm. By monitoring participants' eye movements while they prepare and produce sentences, we expect to have a measure of which information is under preparation from the moment the stimulus is presented until speech starts (see Griffin, 2004, and Konopka & Brown-Schmidt, 2014, for reviews of the uses of the eye-tracking methodology in language production). This method, along with cross-linguistic differences between languages, enables to examine the interaction of different planning processes.

As was introduced in chapter 1, and reproduced in Figure 4.1, Spanish RCs are head-initial structures. In other words, when producing an RC in Spanish, the HN will be uttered in first place, regardless of its grammatical function within the subordinate clause. Spanish object RCs (unlike subject RCs, which follow the canonical SVO word order) allow for greater word order flexibility inside the subordinate clause than Japanese, but the preferred order places the subject after the verb, thus resulting in no differences in surface word order between subject and object RCs. As can be seen in Figure 4.1, the order HN – verb – subordinate NP is kept constant across subject and object-RCs.

(1) **Spanish:**

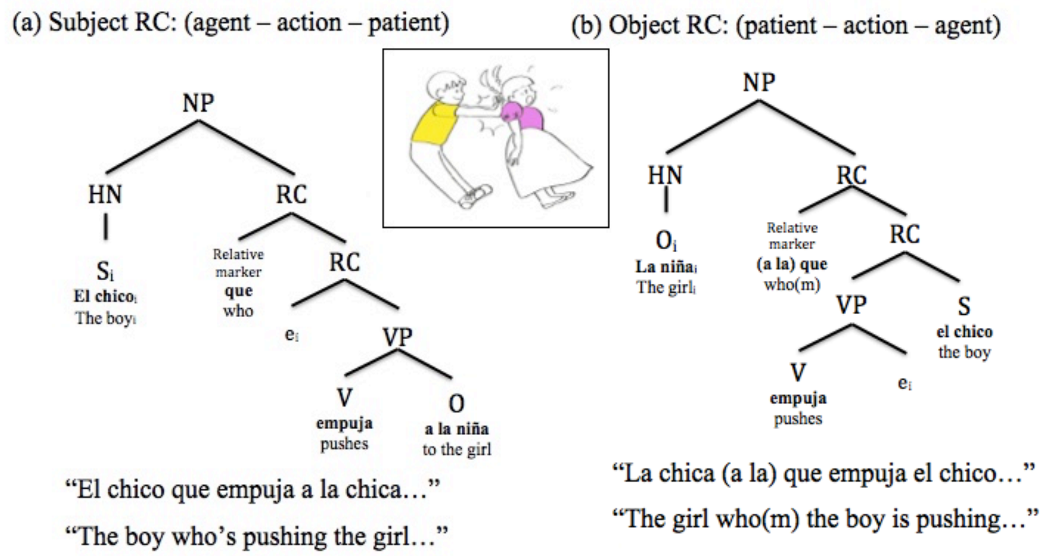


Figure 4.1. Spanish subject RCs (a) and object RCs (b) syntactic and thematic role relations and constituent order.

In contrast, Japanese, a head-final language, places the HN of the RC after the subordinate clause. Thus, interestingly, in Japanese the syntactically highest element of the RC is not the first placed element, but the last one, as can be seen in Figure 4.2, which yields the opposite word order than in Spanish. As a result, word order between Japanese and Spanish RCs is completely the opposite.

(2) **Japanese:**

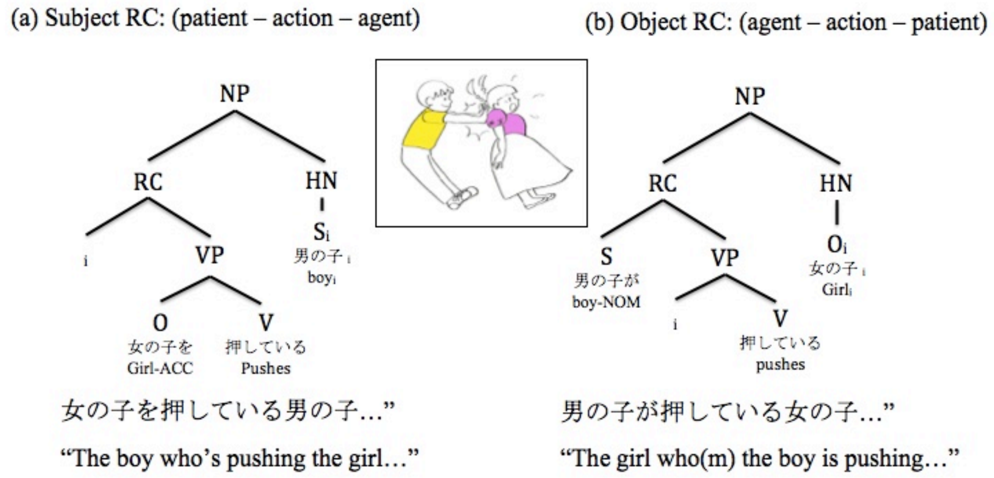


Figure 4.2. Japanese subject RCs (a) and object RCs (b) syntactic and thematic role relations and constituent order.

The comparison of the same structure in two languages with opposite word order will allow to figure out whether the prioritized information when undertaking syntactic planning is relational or non-relational, as we argue in the next section.

Additionally, the animacy of both agent and patient was manipulated in our studies, thereby controlling the effects of conceptual saliency of individual items and allowing for an analysis of its influence from the very beginning of speech planning.

Finally, in RCs with the patient as the HN, we compare sentences expressed in the active voice to those uttered in the passive voice. As can be seen in (3) below, word order remains identical in both structures within each language, whilst grammatical functions differ between them.

(3) RCs with patient as the HN:

a. Spanish object RCs (with the patient as HN) in active voice:

La niña (a la) que empuja el chico lleva un vestido rosa

The girl (to) who pushes the boy wears a dress pink

"The girl who(m) the boy is pushing is wearing a pink dress"

(Word Order: Patient / Object (HN) – Verb – Agent / Subject)

b. Spanish subject RCs (with the patient as HN) in passive voice:

La niña que es empujada por el chico lleva un vestido rosa

The girl who is pushed by the boy wears a dress pink

“The girl who is pushed by the boy is wearing a pink dress”

(Word Order: Patient / Subject (HN) – Verb – Agent / by-phrase)

c. Japanese object RCs (with the patient as HN) in active voice:

Otoko-no-ko-ga oshiteiru onna-no-ko-wa pinku-no doresu-o kiiteimasu

Boy-NOM pushes girl-TOP pink dress-ACC wears

“The girl who(m) the boy pushes wears a pink dress”

(Word Order: Agent / Subject– Verb – Patient / Object (HN))

d. Japanese subject RCs (with the patient as HN) in passive voice:

Otoko-no-ko-ni osareteiru onna-no-ko-wa pinku-no doresu-o kiiteimasu

Boy-DAT pushes girl-TOP pink dress-ACC wears

“The girl who is pushed by the boy wears a pink dress”

(Word Order: Agent / by-phrase – Verb – Patient / Subject (HN))

Thus, the comparison of sentences with agent vs. patient as HN between languages (examples 1 and 2), on the one hand, and of passive vs. active sentences with the patient as HN within a given language (example 3), on the other, will provide an opportunity to assess the relative importance of structural and lexical information in the planning process, enabling us to test the tenets of linear (e.g. Gleitman *et al*., 2007; Brown-Schmidt & Konopka, 2008; Myachikov & Tomlin, 2008); and hierarchical incrementality accounts (e.g. Griffin & Bock, 2000; Van de Velde *et al*., 2014; Bock *et al*, 2003; Norcliffe *et al*., 2015) against each other.

Bearing in mind the interesting contrasts between Spanish and Japanese RC structures, our studies will focus on three different periods of time during sentence preparation and production, each hypothetically designated to address a different issue regarding speech planning: (1) an initial period during which the speaker creates a conceptual representation of the event; (2) the period following conceptualization, where the conceptual representation of the message is delivered for grammatical encoding; and (3) the stage at which linearization and lexical insertion are supposed to take place. The analysis of these periods will allow us to investigate the relation between conceptualization and grammatical encoding, and between grammatical encoding and linearization, while analyzing which type of information, relational or non-relational, is prioritized when encoding an utterance. We will address each of these issues in turn when presenting our hypotheses for each of the temporal windows.

1. Apprehension period: how is the starting point chosen?

The main question concerning this first period of analysis is whether the saliency of individual lexical items determines the starting point in sentence planning, thus guiding word order planning in complex structures. In order to explore this, we manipulated the animacy of the agent and the patient, with three different combinations in both RCs with the agent and the patient as HN:

(4) a. Animate Agent– Animate Patient (AA):

- El chico que empuja a la chica
- 少女を押している少年 (Girl-ACC is pushing boy)

  "The boy who pushes the girl"
- La chica a la que empuja el chico
- 少年が押している少女 (boy-NOM is pushing girl)

  "The girl whom the boy is pushing".

b. Animate Agent– Inanimate Patient (AI):

- El campesino que empuja el carro
- 一輪車を押している農人 (cart-ACC is pushing farmer)

  "The farmer who is pushing the cart"
- El carro que empuja el campesino
- 農人が押している一輪車 (farmer-NOM is pushing cart)

  "The cart that the farmer is pushing"

c. Inanimate Agent– Animate Patient (IA)[37]:

- El camión que empuja al policía
- 警察官を押しているトラック (policeman-ACC is pushing truck)

  "The truck that is pushing the policeman"
- El policía al que empuja el camión
- トラックが押している警察官 (truck-NOM is pushing policeman)

  "The policeman whom the truck is pushing"

First, we will analyze the effects of animacy on spoken responses in both languages. As it was reviewed in the previous chapter, animate HNs yield an increased proportion of passive sentences in RCs regardless of word order (Spanish: Gennari *et al.*, 2012; Japanese: Montag & MacDonald, 2009). For that reason, we expect to find a similar pattern in participants' responses, including the novel condition IA.

We will subsequently compare gaze patterns to agents and patients under various animacy combinations during the first 400 ms. from picture onset (which is claimed to be the average time window needed to complete the apprehension of the scene (Griffin & Bock, 2000)). We will analyze, both in Spanish and Japanese, (1) whether animate elements attract more gazes from picture onset; and (2) whether there exists a relation between the direction of gazes and the likelihood to promote the animate element to the subject function. If there is an increased proportion of gazes directed to animate elements, and this shows a relation with the final shape of the utterance, it may be concluded that the conceptual saliency of isolated elements does have an effect on linguistic encoding. Importantly, word order is not free in RCs, either in Spanish or in Japanese. The analysis of the effects of animacy on voice will only focus on RCs in which the HN is the patient, as these are the only ones that allow

[37] The Inanimate-Animate condition was included in order to test the effect of semantic similarity of nouns (Gennari *et al.*, 2012) on gazes and speech patterns. We are aware that inanimate elements cannot be "agents" as such, as they lack agentivity. They carry the role of doer, in opposition to the undergoer, in our scenes. However, for sake of uniformity with previous studies using conceptual accessibility, we will keep the nomenclature "agent" for the doer and "patient" for the undergoer. The Inanimate-Animate condition, due to its lower degree of familiarity, will help reveal the potential effects of scene prototypicality in general, along with the saliency of the isolated elements, regardless of their thematic role.

variation due to voice. In Spanish, RCs with the patient as HN, animate HNs are expected to yield a higher proportion of passive sentences, and the same holds true for Japanese (replicating previous studies – see chapter 2), as a consequence of promoting the animate element to the subject position in the sentence, but with no variation in word order. In Spanish, animate HNs that bear the subject function are also placed at initial position. The trends to assign the subject function and/or the initial position to the animate element, and also to the syntactically most dominant one (the HN in this case) converge to yield these results, thus helping the mapping of the various elements onto the ongoing planning of the utterance. In other words, all cues match in the case of Spanish: if the animate (conceptually more salient) element happens to be the HN (the syntactically dominant element), it has to be placed at the beginning of the sentence (the linearly most salient position), and thus is more likely to be assigned the subject function (the most salient grammatical function).

In contrast, in Japanese, promoting animate HNs to the subject function will result in an utterance in which the animate element, also the subject and the HN of the sentence, is placed at the end. Despite this, we expect to find an increase of passive sentences when the HN is an animate entity and bears the patient role, replicating Montag & Macdonald's (2009) results. An analysis of gaze patterns and speech latencies between conditions will shed light on the mechanisms underlying the promotion of the animate HN to the subject function.

It is difficult to fit the tenets of a strong version of linear incrementality models with the results reported by Montag & MacDonald (2009. However, it might be the case that more accessible elements are first assigned grammatical function, regardless of the final word order. Note that the difference between structural and linear incrementality models in this regard relies on the contrast between planning the whole clause vs. planning one element at a time, beginning with the element that will be assigned grammatical function and/or linear order in the first place. In this case, however, the grammatical function to the animate element will be assigned first, but this element will have to be placed at the end of the clause. Due to the strict word order of RCs, it is reasonable to assume that this is indeed what happens. Thus, if we expect that conceptual accessibility of single elements facilitates an earlier encoding of such elements, placing the animate noun at the end of the sentence, as would be the case in Japanese, would result in a more costly process, since speakers would have to buffer the first encoded element until the end of the sentence. If so, the promotion of the HN to the subject position would likely help alleviate the buffering costs (by making this element more salient). On the other hand, if the assignment of grammatical functions is guided by a more general plan, as structural incrementality predicts, the animate element will be assigned the subject function, given its prototypicality, regardless of its linear position in the final utterance, either initial in Spanish or final in Japanese. However, in the case of Japanese, the lemma of the animate noun (the HN placed at final position) will not be activated yet, and so the speaker will not have to buffer it until the end of the sentence. Both linear and structural incremental planning will result in the same kind of utterance, but the differences will be measurable by means of gaze movements and their relation with the final output.

In this regard, if the lemma of the animate noun is the first to be accessed and planned, even if placed at the end –as is the case in Japanese–, we expect that (1) animate elements will be significantly more focused than inanimate elements *before* 400 ms., and those trials in which the animate element is fixated more extensively will result in passive sentences to a greater extent. This is bound to happen both in Spanish and Japanese. (2) Specifically for Japanese, there will be longer speech onset latencies in RCs with the patient as HN when the HN is animate (i.e. the AA and IA conditions) compared to cases where the HN is inanimate (i.e. the AI condition), as speakers will have to buffer the animate element that they processed at the very beginning until the end of the sentence. In the case of Spanish, the opposite would hold true, as the first fixated element will be the one that is placed earlier, thus maximizing incremental planning. Finally (3) we expect that, should conceptual accessibility lead to a speech planning based on the activation of a single lemma first, in Japanese, participants will have to go back to focus the animate HN more often than the inanimate HN, thus reducing the differences in gazes between them. These differences should not be visible in Spanish.

On the other hand, if speakers first create a structural scaffold to be filled with the corresponding lemmas, we expect that (1) there will not be a significant relation between early fixations (i.e. before 400 ms.) to the animate HNs and the proportion of passive sentences, in both Spanish and Japanese. In other words, we expect there will be no differences in early fixations to the animate HN between RCs uttered in active and in passive voice. (2) Additionally, we expect that speech onset latencies in AA and IA sentences will not be longer than in AI sentences in general, both in Spanish and Japanese[38]. (3) Finally, we expect that there will be no differences in the proportion of gazes to animate and inanimate elements between Spanish and Japanese, as the structure will have been created beforehand, leading to a similar process in both languages. As a result, once lexical retrieval has started there will be no differences due to the animacy of the HN. Note that, as we argue later (see point 3 below), there could be differences between active and passive sentences *among* different animacy combinations. Any interaction effects will reveal an influence of the conceptual accessibility of isolated elements. We will explain this point in more detail in point 3 below.

2. From apprehension to lexical retrieval: prioritized information in linguistic encoding

The second period of analysis we will take into account is concerned with the information that is encoded *after* apprehension has finished, that is, right *after* 400 ms. This analysis will allow us to investigate what kind of information, either relational or non-relational, is prioritized in speech planning; in other words, whether incrementality is hierarchically or linearly driven. As it was reviewed in previous chapter, previous studies have suggested that speakers take less than 400 ms. (cf. Griffin & Bock, 2000) to understand the scene they are presented with and decide how they are going to plan their utterance. From this moment, they start planning the lexical items and fixate each of the elements of the sentence incrementally, by means of what has been called "name-related gazes" (gazes directed to the elements in order to retrieve their lexical forms (Griffin, 2004)). In chapter 2, several studies making use of eye tracking were reviewed, but in all of them the first element was at the same time the most prominent in the sentence. In those studies, speakers started focusing on the subject, the first element, right after apprehension. However, it is not clear

[38] It is important to note, however, that in AA sentences lexical competence might come into play (Gennari *et al.*, 2012), affecting speech onset latency in this kind of sentences, which can result in longer delays. Similarly, the IA condition might be confounding due to the low prototypicality of the whole scene. Both the effects of lexical competence and scene prototypicality will equally affect whether there is lexical or hierarchical incrementality. For that reason, we will still be able to observe differences beyond those due to prioritizing the planning of the animate element (linear incrementality) over the construction of a whole scene (structural incrementality).

whether they were at the same time preparing a whole scaffold of the sentence (structurally-driven incrementality), or if this structure was planned after the lemma of the subject was accessed (lexically-driven incrementality). An appropriate comparison should be drawn with a language in which the most dominant element of the sentence is not the first element, thus providing an opportunity to disentangle structural from lexical accounts. This issue will be examined in the current study by comparing RC planning *between* languages: Spanish RCs and Japanese RCs exhibit the opposite word order, but despite this, the internal structure of the RC remains identical, allowing for comparison of structural and linear planning.

In Spanish, the HN is always clause-initial, thus yielding the problem that was mentioned above: in a construction like RCs it is virtually impossible to tell apart the most dominant position from the first uttered element. However, Japanese shows the opposite pattern, which enables to identify the kind of information that is prioritized in order to undertake linguistic encoding: either relational, where the construction of an overall structure takes priority, or non-relational, where the retrieval of lexical items as they are going to be uttered determines the choice of sentence structure.

If linear incrementality takes the lead, with stronger reliance on non-relational information, we would expect to find a pattern in which the items are fixated in the same order as they are uttered. Critically, we would expect this to be the case for both Spanish and Japanese, regardless of the position of the HN. As a consequence, the order of gazes would be reversed between Japanese and Spanish in subject and object RCs. Gazes after apprehension will be guided to the first placed element, and only after that the grammatical relations between this element and the following will be decided.

On the other hand, if speech planning undergoes hierarchical incrementality, thus favoring relational information, we would expect a pattern in which, following apprehension of the event, the structure is planned before name-related gazes take place. This pattern would be visible both in Spanish and Japanese. Importantly, in contrast with simple transitive clauses, where planning of the structure results in convergent gazes between agent and patient (Gleitman *et al.,* 2007, Konopka & Meyer, 2014), we expect that the construction of the structure in RCs should produce a pattern of increased gazes to the HN, as the most dominant element, on which the RC is dependent. As a result, if linguistic encoding is guided hierarchically, participants should start focusing on the HN within 400 ms. after picture onset, and these gazes would last until lexical retrieval (name-related gazes) starts. From that moment, Spanish speakers will keep focusing the HN, as the first uttered element, while Japanese speakers are expected to switch their gazes, turning to the first uttered element before speech onset occurs. A summary of predictions is shown in Table 4.1.

	Subject RC	Object RC
If linear incrementality takes place	Japanese ≠ Spanish	Japanese ≠ Spanish
If hierarchical incrementality takes place	Japanese = Spanish	Japanese = Spanish

Table 4.1. Summary of expected differences in the order of gazes from 400 ms. onwards, before name-related gazes start: Constant grammatical function and different word order across languages.

In this case, hypotheses are proposed concerning the comparison between Spanish and Japanese. However, as it was mentioned above, in Spanish prioritization of relational and non-relational information is confounded. We are aware that due to cross-linguistic differences in the way speakers of both languages handle relational and non-relational processes, we cannot extrapolate directly our findings in Japanese to Spanish, with the possibility arising that one language relies more on linear incrementality and the other focuses more on global structural relations between elements. We will analyze, nevertheless, the similarities and differences between both languages in order to have a clearer idea of the main tenets that guide RC planning in a head initial language, such as Spanish.

3. Speech planning during lexical retrieval: Is there an interaction between grammatical function assignment and linearization during this stage?

Finally, in order to explore whether there is an overlap between function assignment and constituent assembly once lexical selection has started, we will analyze differences in planning processes dependent on the voice of speakers' responses *within* each language. At this point of time, our aim is to analyze whether function assignment is fully established before initiating lexical retrieval, or there are points in which both processes overlap and can be independently measured. The same holds true if a prioritization of non-relational processes were found at the previous stage.

To explore this issue, we will focus here only on sentences with the patient as HN, in both languages. As mentioned above, RCs with the patient as HN in active and in passive voice show an identical word order in both cases, but not the grammatical roles assigned to the participants in the event.

As can be seen in example (3) above, and reproduced here for clarity, in Spanish, the selection of an object RC in this type of utterances will result in a structure in which the first uttered element is not the subject, but the object of the sentence (example 3a), reversing the order found in simple transitive clauses. This brings about a situation in which the syntactically most prominent element in the RC (the HN) is sentence initial but does not hold the most prominent grammatical function within the relative clause, whilst the constituent bearing the subject function is relegated to a subordinate role.

(3') RCs with patient as the HN:

a. Spanish object RCs (with the patient as HN) in active voice:

   La niña (a la) que empuja el chico lleva un vestido rosa

   The girl (to) who pushes the boy wears a dress pink

   "The girl who(m) the boy is pushing is wearing a pink dress"

   (Word Order: Patient / Object (HN) – Verb – Agent / Subject)

b. Spanish subject RCs (with the patient as HN) in passive voice:

   La niña que es empujada por el chico lleva un vestido rosa

   The girl who is pushed by the boy wears a dress pink

   "The girl who is pushed by the boy is wearing a pink dress"

   (Word Order: Patient / Subject (HN) – Verb – Agent / by-phrase)

c. Japanese object RCs (with the patient as HN) in active voice:

   Otoko-no-ko-ga oshiteiru onna-no-ko-wa pinku-no doresu-o kiiteimasu

   Boy-NOM pushes girl-TOP pink dress-ACC wears

   "The girl who(m) the boy pushes wears a pink dress"

(Word Order: Agent / Subject– Verb – Patient / Object (HN))

d. Japanese subject RCs (with the patient as HN) in passive voice:

Otoko-no-ko-ni osareteiru onna-no-ko-wa pinku-no doresu-o kiiteimasu

Boy-DAT pushes girl-TOP pink dress-ACC wears

"The girl who is pushed by the boy wears a pink dress"

(Word Order: Agent / by-phrase – Verb – Patient / Subject (HN))

A similar pattern is found in Japanese. Here, object RCs (see example 3c) result in a sentence in which the subject is the first uttered element, though not the syntactically most prominent one, which appears at the end of the sentence, for Japanese is a head-final language.

A comparison between passive and active responses within each language will provide a way to find out if there are differences in the assignment of grammatical functions while keeping word order constant. It will provide evidence of any differences due to grammatical function assignment after name-related gazes (Griffin, 2004) have been set off, regardless of whether structural choices have been previously established (see issue number 2 above).

Under a strong version of *linear incrementality*, grammatical function assignment is inseparable from the retrieval of the corresponding lemmas that carry such functions (e.g. Gleitman *et al.,* 2007; see also Bock *et al.,* 2004). As a consequence, it can be hypothesized that in this case, the NPs denoting each of the participants in the event, along with their corresponding grammatical functions, will be retrieved in the same order as they will be uttered, with no visible differences between active and passive sentences. On the other hand, under a strong version of *hierarchical incrementality*, grammatical functions should be completely established before lexical retrieval begins to operate in a serial fashion. Thus, similarly to the previous case, there should be no visible differences in the retrieval of lexical elements as they are going to be produced, since grammatical functions are *fully assigned* before lexical retrieval starts.

In contrast, if grammatical functions are *not fully assigned* before constituent assembly begins, we should find differences in planning processes between passive subject RCs and active object RCs (as summarized in Table 4.2). These differences should be visible regardless of the information that was prioritized at the previous stage of grammatical encoding from conceptualization.

	Spanish patient-HN RC	Japanese patient-HN RC
No overlap between stages	Active = Passive	Active = Passive
Overlap between stages	Active ≠ Passive	Active ≠ Passive

Table 4.2. Summary of expected differences in gaze order once name-related gazes have started: Constant word order and different grammatical function across languages.

In what follows, we present two experiments intended to clarify the time-course of the production of RCs, focusing on the three critical periods discussed above. We used eye-tracking methodology, monitoring participants' eye movements before and during the production of RCs in a visual-world paradigm. In the first experiment, we report data from Spanish. The second experiment was conducted in Japanese, using the same method and procedure. Subsequently, a comparison between both experiments will be carried out. The analysis of each experiment in isolation will allow us to look into the intralinguistic processes, helping us understand mainly the first and third issues described above. The comparison of both experiments will help us understand the role of word order differences after scene apprehension is over.

Experiment 1: Relative Clause production in Spanish

Method

Participants.

Thirty-one Spanish native speakers participated in this study. They were undergraduate or graduate students at the Universidad Autónoma de Madrid. Mean age was 22.68, with a range of 17 to 42. There were 23 women and 8 men. All participants reported normal or corrected-to-normal vision.

Materials.

Colored pictures were presented for description: 30 critical and 30 filler pictures. Each critical picture depicted four participants: two of them were involved in a transitive action, while the other two remained inactive. Examples of critical pictures can be seen in Figure 4.3. Position of the four elements was counterbalanced in the up-down and left-right axes, so that every element could be found with the agent in the left or in the right, up or down, and accordingly, the same holds true for the patient. There were ten different actions, each coupled with a different animacy condition (see (4) above).

Similarly, fillers consisted of a four-participant scene showing either intransitive actions or objects with contrasting sizes (in such cases, participants had to answer with the size-contrasting word: e.g. "(It is) the small ball". Each experimental item was presented twice during the task, with questions referring either to the agent or to the patient of the event.

The full list of experimental and filler pictures coupled with their corresponding verbs and questions can be found in Appendices 1 and 2, respectively.

(a) Animate-Animate (b) Animate-Inanimate (c) Inanimate-Animate

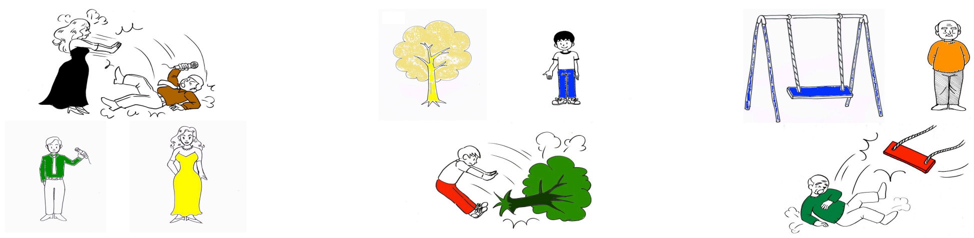


Figure 4.3. Examples of critical items in the three animacy conditions (action: knock down).

Apparatus

In order to measure participants' eye movements to the different elements, we used Tobii T120, with a sampling rate of 60Hz. Stimuli presentation and data collection were performed using Tobii Studio 3.2. Verbal responses were recorded and transcribed for analysis.

Procedure

Participants were tested individually. Before the task, participants underwent a training session in which they examined separately all the animate and inanimate participants and the main actions that would be presented during the task (the pictures used in the study can be found in Appendix 1 and Appendix 2). They were asked to name them, and when they did not come up with a response, the experimenter provided the correct answer. This was made to ensure participants understood all the elements involved in the scenes, and the actions represented therein. Despite providing them with a correct answer, they were encouraged to respond with the most natural description during the task and not think too long about the name that every element should receive.

Participants were seated in front of a screen-based Tobii T120, at a distance of approx. 70 cm., in a fixed chair, with their eyes fixated on the center of the screen. They were requested to sit comfortably, in a position they could hold as long as the experiment lasted.

After the training session, the experimental task began. First, the task was explained to each participant and a built-in 9-point calibration was conducted. There were four practice items before the task. The task consisted of answering a question presented on the screen, referring to the agent or to the patient of the event. These questions asked about the color or shape of one of the participants involved in the transitive action of the picture (e.g. "¿Quién lleva un vestido negro?", "Who is wearing a black dress?"). This screen appeared for 4 seconds. After that, the following screen presented a verb in the infinitive form (e.g. "Derribar" – "To knock down"), during 3 seconds. Participants were instructed to use this verb in their answers, in any conjugation and form they preferred. Finally, the picture appeared on the screen. At this point, the participant had to answer the question while using the provided verb. When they finished the sentence, they pressed the space key in order to move forward to the next item. There was no time pressure, but participants were encouraged to answer with the first and most natural response that came to mind.

Results

Pattern of spoken responses

Data analysis. Produced sentences were transcribed and analyzed. Responses that did not correspond to RCs or differed in meaning or HN were excluded from the analysis. On the other hand, responses that made use of a different verb, but that kept the same meaning and transitive structure were included in the analysis. Responses in which the subordinate noun was excluded (e.g. "El chico que es empujado" – "The boy who is being pushed”) were included in the analysis of spoken responses. Excluded responses amounted to 19.03% of all responses (i.e. 347 out of 1823 total responses).

Responses were classified as active, passive or impersonal sentences. The ‘impersonal sentences’ category contains both impersonal and reflexive sentences[39].

We analyzed the proportion of passive sentences with a linear mixed model analysis, with RC type (Agent HN / Patient HN) and Animacy combination (AA / AI / IA) as fixed factors and crossed-random slopes for subjects and items. The full list of proportions and statistical values can be found in Appendix 3.

Results. As expected, sentences with the agent as HN were produced as actives in almost all cases. On the other hand, items with the patient as HN yielded a wider variety of responses, depending on the animacy of the elements involved in the picture. There was a main effect of both RC type ($p < 0.001$) and animacy, and a significant interaction between them ($p < 0.05$). When the agent is the HN, sentences are produced almost exclusively in active form, with the exception of a 5.4% of impersonal sentences in the IA condition. When the patient is the HN, there are a wider variety of responses, depending on the animacy of the elements. AA and IA sentences did not differ from each other, both producing a greater number of passive sentences than items with AI ($p < 0.001$). Impersonal sentences were present in all three animacy combinations, with a higher percentage in the IA condition (13.61%), although the difference with the other two animacy combinations was not significant (Figure 4.4). That is, animate patients were more prone to be produced as subjects than inanimate ones.

[39] As it was explained in Chapter 2 with reference to Gennari *et al.* (2012) study, Spanish allows sentences in which the subject is not overtly expressed, while the object remains in place. In these sentences, the agent is implicitly referred to by a plural inflection in the verb (e.g. “El soldado al que están empujando” – The soldier whom (they) are pushing).

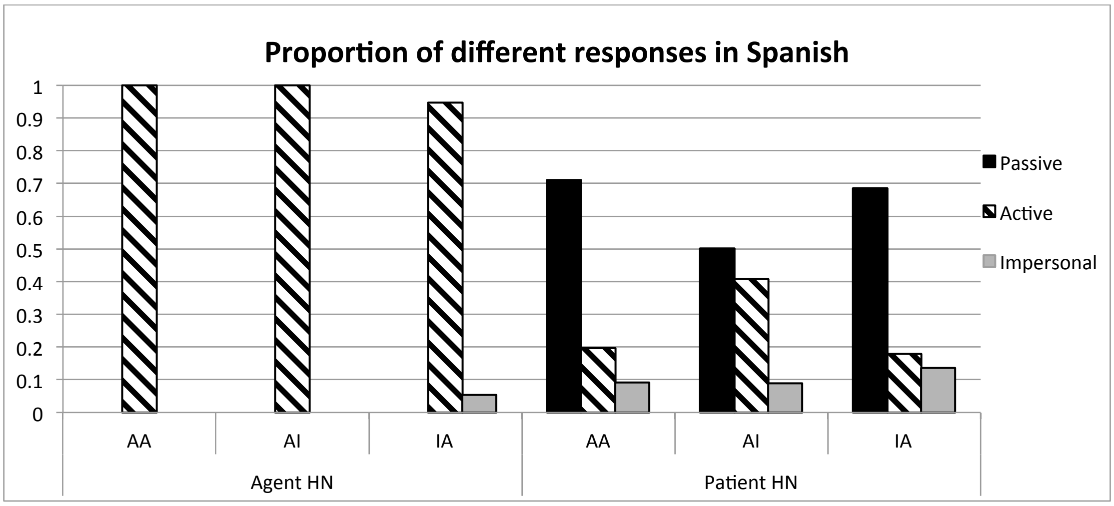


Figure 4.4. Proportion of active, passive and impersonal responses in Spanish across the different animacy combinations in the Agent-HN condition (left) and the Patient-HN condition (right).

<u>*Pattern of gazes*</u>

Data analysis. Furthermore, we analyzed fixations to the agent and the patient of the actions from picture onset up to 6000 ms. In order to ensure that responses were constant enough to allow comparison, we followed restrictive exclusion criteria. Accordingly, we excluded from gaze analysis the following types of responses:

a. Responses that did not start with the relative clause. In this type we included those responses that started with the main clause instead of the relative clause (e.g. “Lleva una camiseta roja la chica que empuja al chico: – Wears a red T-shirt the girl who pushes the boy). This type of responses resulted in sentences in which the HN was not the first uttered element, thus making them difficult to compare with other responses. Excluded responses due to this cause were 82 out of 1468 valid responses (5.59% of responses).

b. Responses that failed to overtly mention the two participants involved in the event (e.g. sentences with the subordinate element dropped, like “El chico que es empujado” – The boy who is pushed). If one of the elements is not mentioned, participants will likely not focus it in a comparable way to when it is overtly mentioned. Thus, any difference found between sentences would be difficult to understand. Excluded responses due to this cause were 230 out of 1468 valid responses (15.67% of responses).

c. Active Object RCs in which the subject preceded the verb. As we saw before, Spanish active object RCs allow flexibility in the position of the subordinate noun, the agent and subject of the subordinate sentence, relative to the embedded verb. However, the preferred order is the one in which the subject follows the verb (e.g. "El chico al que empuja la chica" – The boy whom pushes the girl), with only a small proportion of sentences being uttered with the subject-verb order (e.g. "El chico al que la chica empuja" – The boy whom the girl pushes). Subsequently, we decided to exclude this last type of responses so as to compare voice-dependent fixation patterns in sentences with minimal changes in word order (thus, keeping the word order HN – verb – agent in both active and passive sentences). The proportion of responses excluded for this reason was only 1 out of 1468 (0.07% of valid responses).

Additionally, Speech Onset (SO) time was measured manually with the Praat software (version 5.3.71) (Boersma & Weenink, 2014). Subsequently, the average and standard deviation of each participant was calculated. Speech latencies greater than 10 seconds and over or below two standard deviations from each participant speech onset average were removed as well. As a result, three participants were excluded, since they did not provide any valid response. Excluded responses under these criteria amounted to 44.87 % of all responses (31.91% of RC responses), with a sample of 28 valid participants, and a total of 1005 responses to analyze[40].

The area of analysis was defined for both agent and patient. Accordingly, proportions of fixations to agent and patient were measured. Gaze points were recorded every 16.67 ms. Subsequently, data points were grouped in 50 ms bins in order to capture the time-course of gazes from picture onset onwards. Figures shown below represent these gaze points.

[40] The 31.91% of responses excluded for gaze pattern analysis were all grammatical and well-formed sentences in Spanish, and thus counted as correct answers for behavioral analysis: they consisted mainly of passive sentences with the by-phrase dropped or sentences that did not begin with the HN (e.g. *Is wearing a pink blouse the girl who pushes the boy*, instead of *The girl who pushes the boy is wearing a pink blouse*). These sentences, however natural and correct in Spanish, would have made it difficult to control the timing of gaze patterns and draw conclusions from them. Thus, despite the high amount of excluded responses resulting from these stringent criteria, we deemed it was the best solution to improve gaze pattern analysis. The same applies for experiment two.

Subsequently, after a first analysis based on general patterns, time windows (TW) larger than 50 ms were defined and used for statistical comparisons. Thus, we defined and statistically compared the following TWs: TW1: 1 – 350 ms., TW2: 350 – 1000 ms., TW3: 1000 – 1800 ms., TW4: 1800 – 2500 ms., TW5: 2500 – 4500 ms, TW6: 4500 – 6000 ms.

We conducted two LME analysis for each TW, with both gazes directed to the agent and gazes directed to the patient as dependent variables. The first one was conducted with all valid sentences, including RC type (agent-HN / patient-HN) and animacy combination (AA / AI / IA) as fixed factors and crossed-random slopes or, when the model was not improved by their inclusion, intercepts for subjects and items. The second analysis was conducted only with RCs with the patient as HN. Included fixed factors here were animacy combination (AA / AI / IA) and voice of the utterance (active / passive).

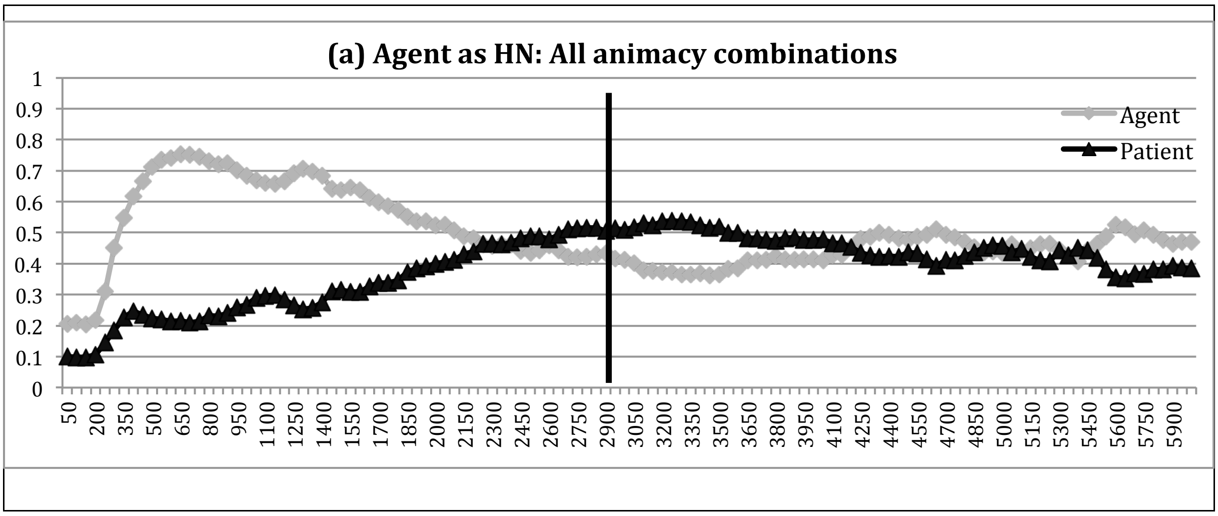
The full list of proportions and statistical values for the two analyses with agent and patient as DV can be found in Appendix 3.

Results. A general view of gaze patterns shows that participants tended to look at the element they were going to mention first in the sentence before speech onset, and that they stopped looking at it before articulating it, with gazes shifted to the next element (Figure 4.5). Speech onset mean is 2984.67 ms., with the Agent-HN sentences having faster SO's than those with Patient-HN ($t = -2.769$, $p = 0.007$). On the other hand, animacy does not seem to play a significant role in speech onset latency ($t = -0.483$), nor the interaction between RC type and animacy. Nevertheless, numerically, IA sentences were the ones that took longer to start, followed closely by AA sentences, AI sentences being the ones with shortest SO (Table 4.3).

	Agent as HN	Patient as HN			General
		Active	Passive	General	
Animate agent - Animate patient	2830	3287.05	3313.27	3306.03	3012.41
Animate agent - Inanimate patient	2856	2959.05	2863.65	2913.36	2879.48
Inanimate agent - Animate patient	2972.54	3123.89	3285.51	3181.75	3065.66
General	2913	3077.71	3183.5	3126.75	2984.67

Table 4.3. Speech onset latencies by Spanish monolinguals of sentences with Agent as HN and Patient as HN in the three animacy combinations.

Regarding issue 1 above, after picture onset there is a short period of about 350 ms. (TW1) in which participants are inspecting the whole scene, in order to locate the element that is being questioned about. Consequently, in TW1 there is no effect of either animacy (AA = 0.21, AI = 0.15, IA = 0.22; t = 0.674; *p* = 0.51) or RC Type, although, in this case, numerically the proportion of gazes to agent and patient when they were the HN was slightly higher than otherwise (gazes to agent: agent-HN = 0.21, patient-HN = 0.17; t = 0.957; *p* = 0.35; gazes to patient: agent-HN = 0.15; patient-HN: 0.11; t = -1.54; *p* = 0.13). This might be due to the fact that gazes start being directed to the HN at around 300 ms, a difference that is reflected in the global mean of gaze proportions in that period of time.



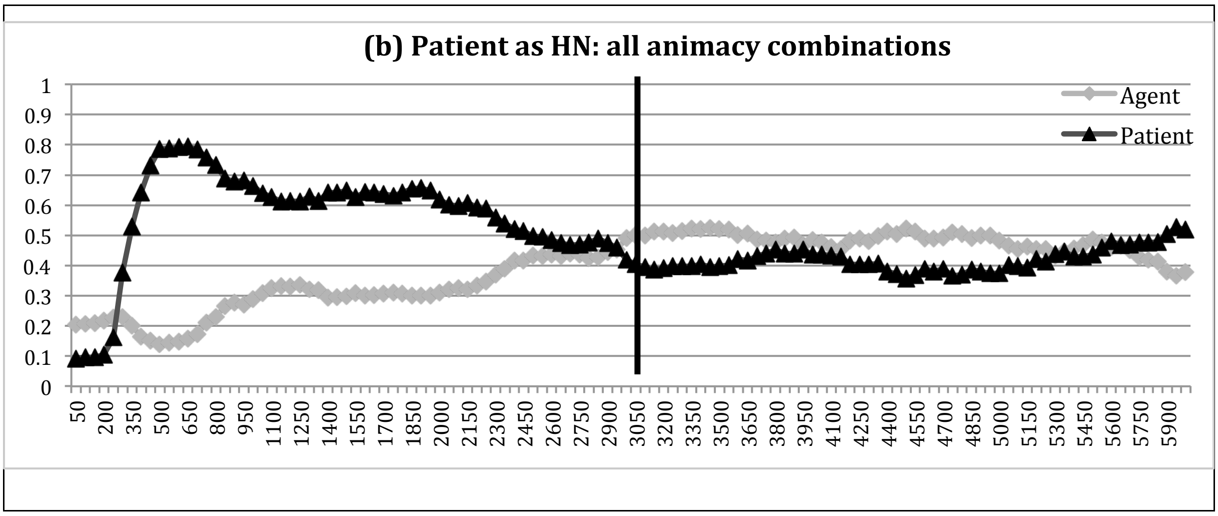
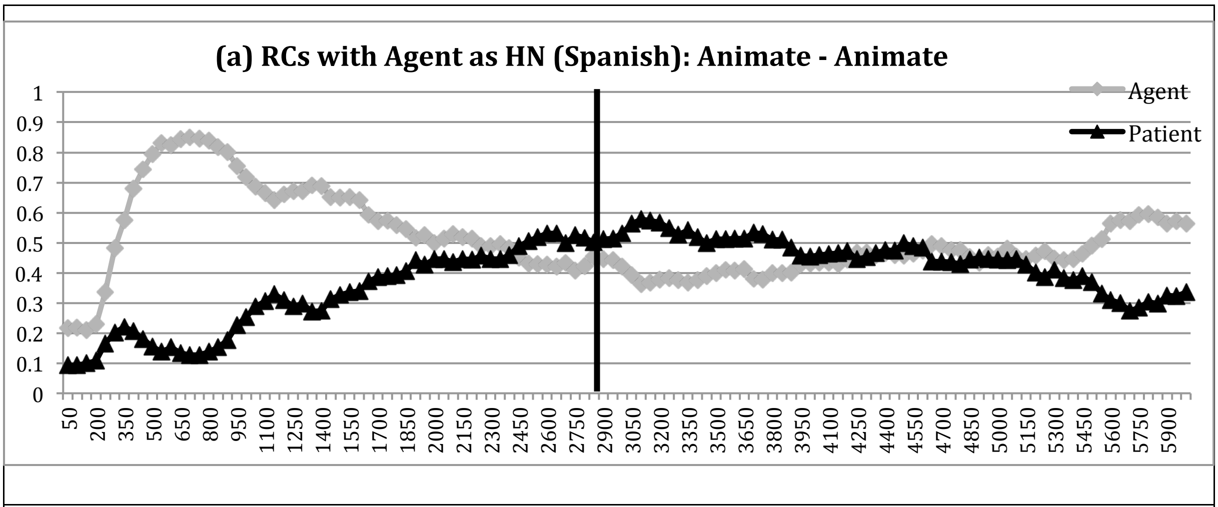
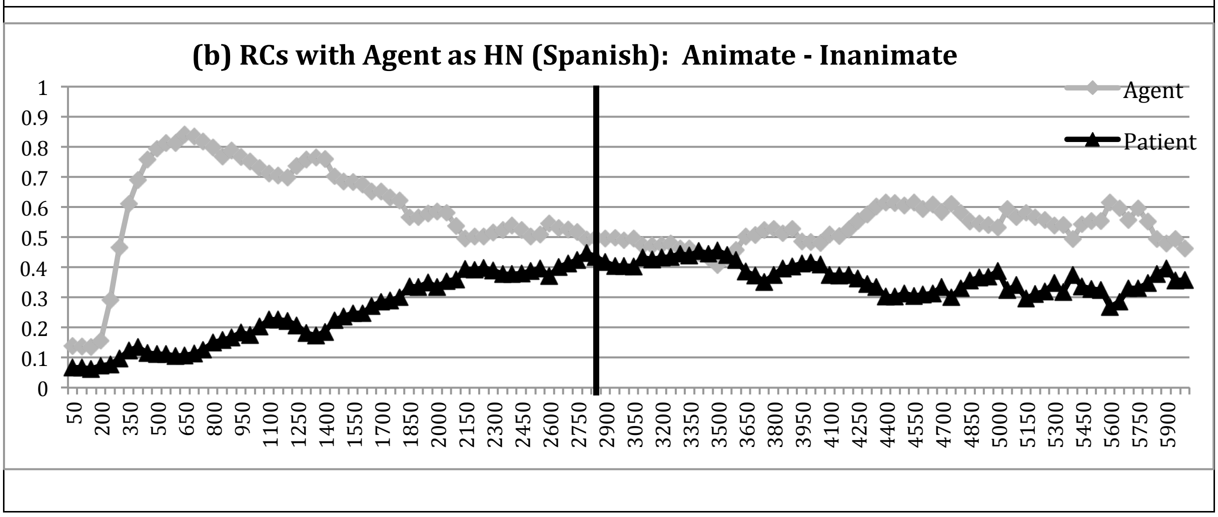


Figure 4.5. General gaze patterns to agent and patient in Spanish RCs with the agent as HN (a) and with the patient as HN (b) (all animacy combinations collapsed): From picture onset until 6000 ms. Vertical line represents speech onset.

From 400 to 2500 ms., (TWs 2 to 4) gaze patterns to both agent and patient differed between RC types (TW2: agent: t = 9.58, $p < 0.001$, patient: t = -10.66, $p < 0.001$; TW3: agent: t = 5.972, $p < 0.001$, patient: t = -7.62, $p < 0.001$; TW4: agent: t = 2.04, $p = 0.047$, patient: t = -3.21, $p = 0.002$), with gazes mostly directed to the HN / first element of the RC (issue 2 above). In contrast, there were no significant differences due to RC type from that moment onwards, even after speech has started.

Regarding animacy, the proportion of gazes to both agent and patient varied with the animacy of the elements, with animate elements attracting more attention. However, this tendency was only significant in TW2 (from 400 to 1000 ms.), for gazes to the agent (t = -2.76, $p = 0.009$), with more gazes directed to the animate element (AI > AA > IA). It reaches significance after SO again, for both gazes to the agent and to the patient in TW6 (from 4500 ms.) (gazes to agent: t= -3.803, $p < 0.001$; gazes to patient: t = 3.133, $p = 0.003$), with the same pattern: more gazes directed to the animate element, that is, the highest rate of fixations to the agent in the AI condition and the highest to the patient in the IA condition. As for the influence of animacy on the form of the final utterance (issue 1), we focused only in sentences with Patient as the HN comparing animacy and voice. Results from this analysis show no significant effects of animacy on TW1, nor an interaction with voice. However, there is a marginally significant main effect of animacy on gazes to the patient (t = 1.819, $p = 0.08$), with more gazes being directed to the patient in IA sentences than in AA and AI sentences (i.e. more gazes directed to the animate element). However, this tendency is not reflected in the final choice of sentence form that participants made, with no differences in the proportion of gazes to patient in IA sentences in active (0.21) and passive sentences (0.20). Figures 4.6 and 4.7 show gazes patterns in both RCs with agent as the HN and RCs with patient as the HN in the three different animacy combinations.





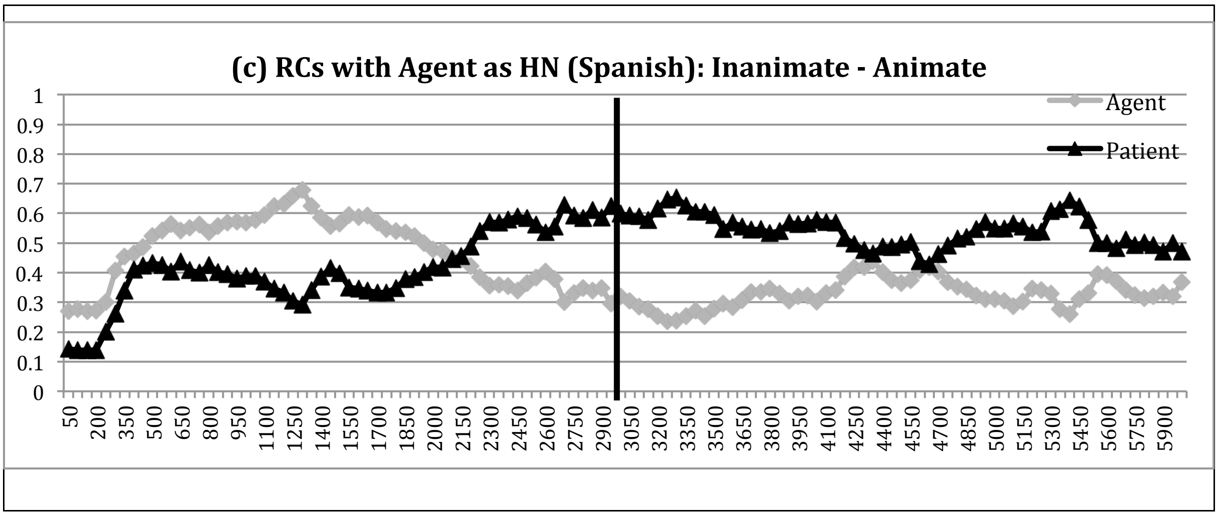
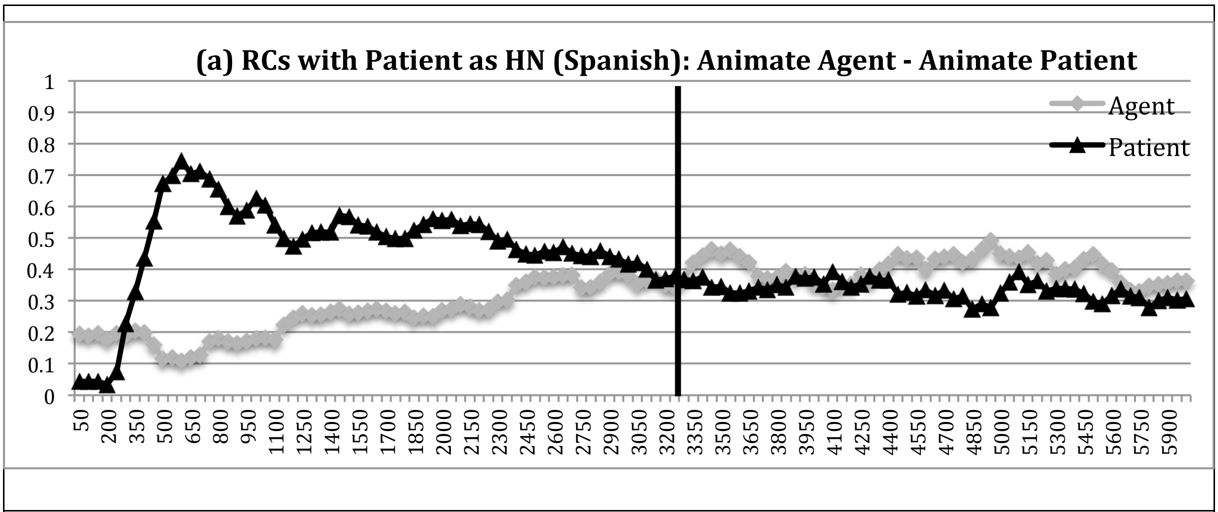
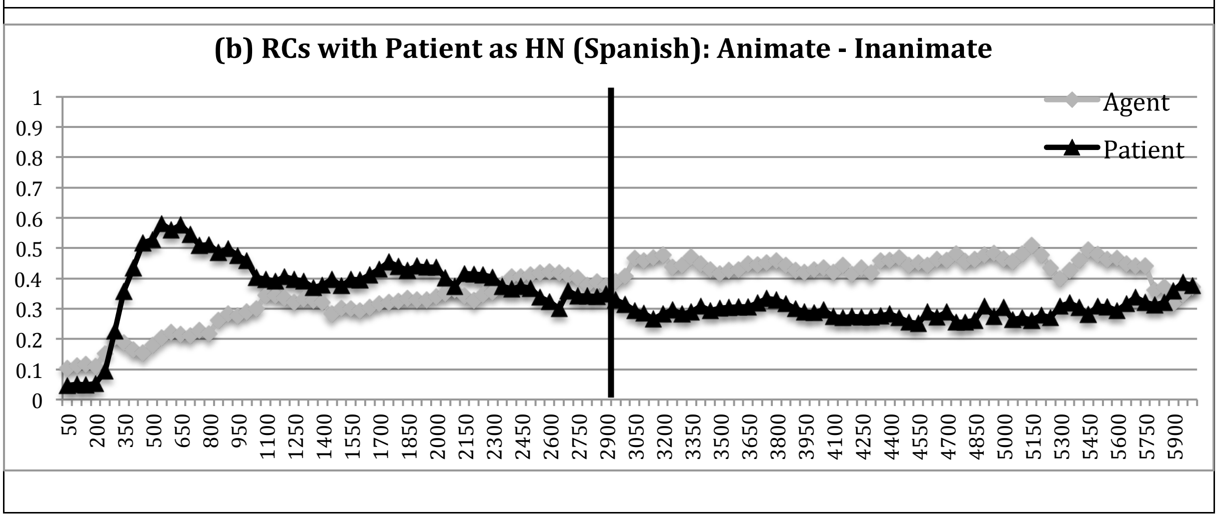


Figure 4.6. Patterns of gazes to agent and patient in Spanish RCs when the agent was the HN, from picture onset until 6000 ms., in all three animacy combinations. Vertical line represents speech onset.





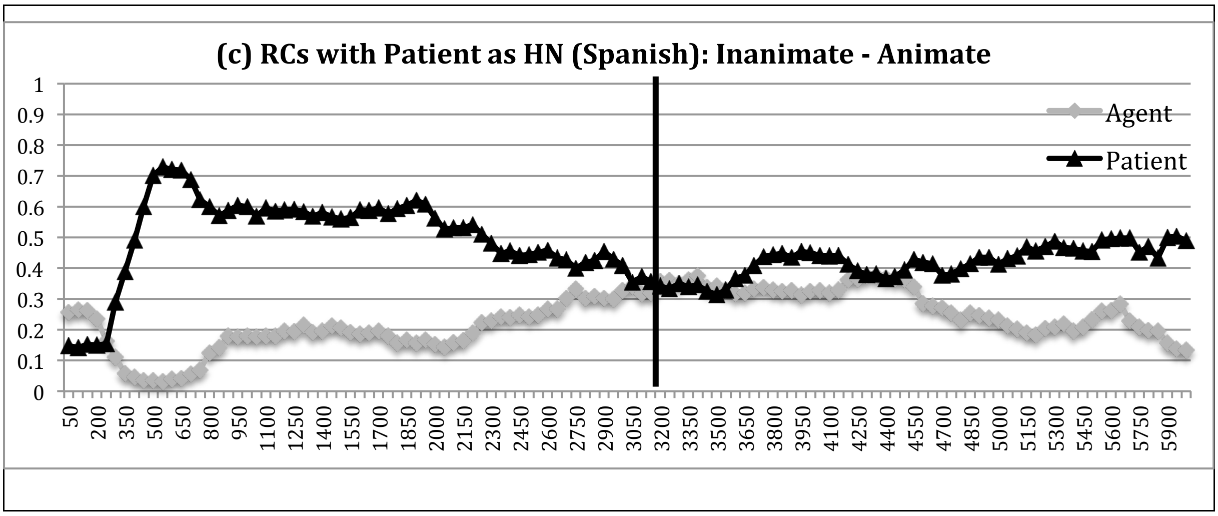
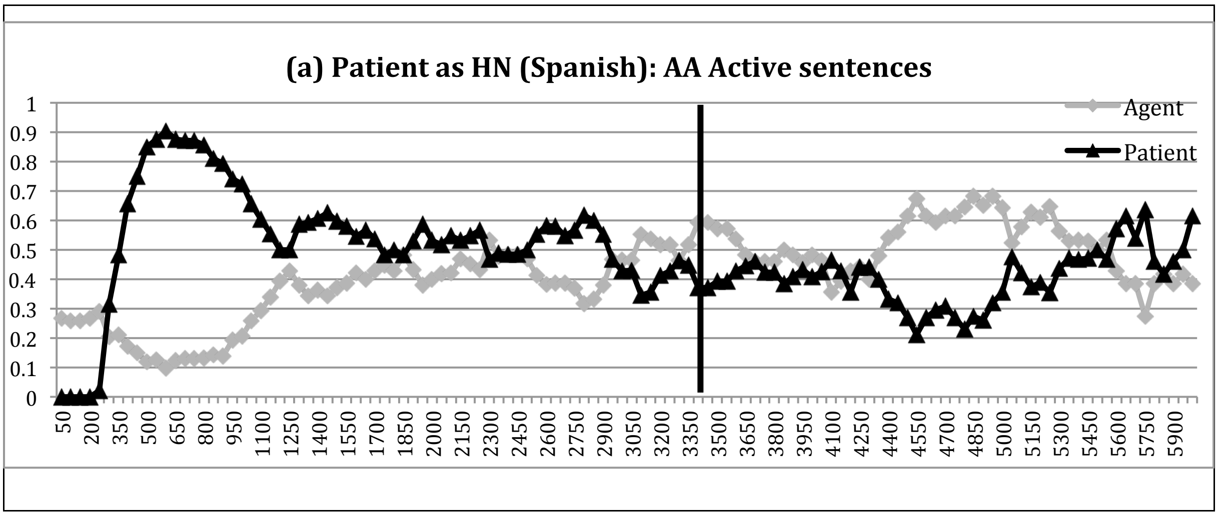
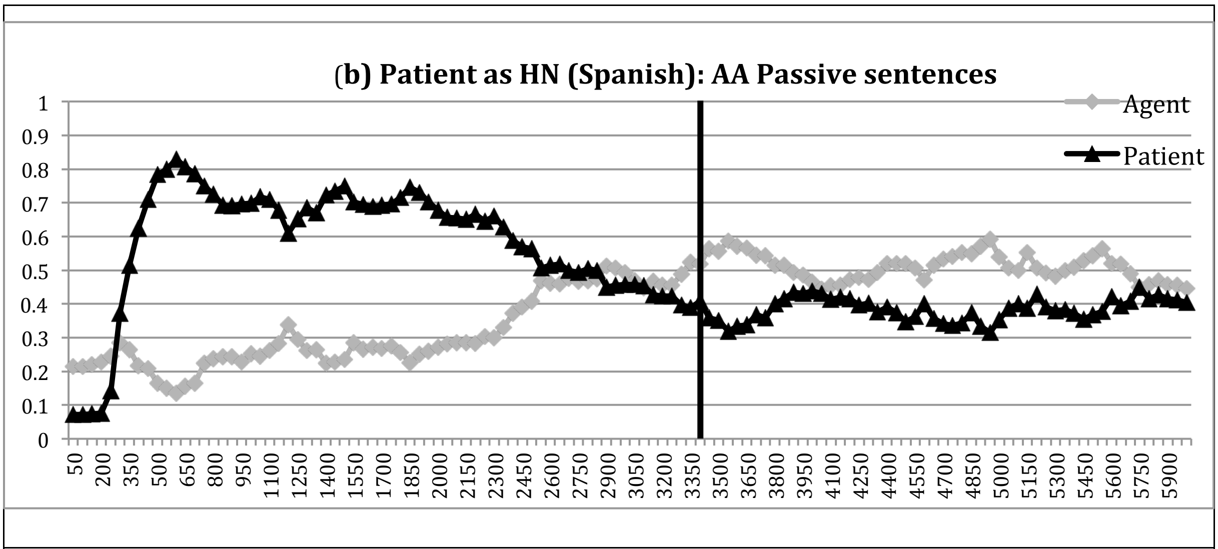
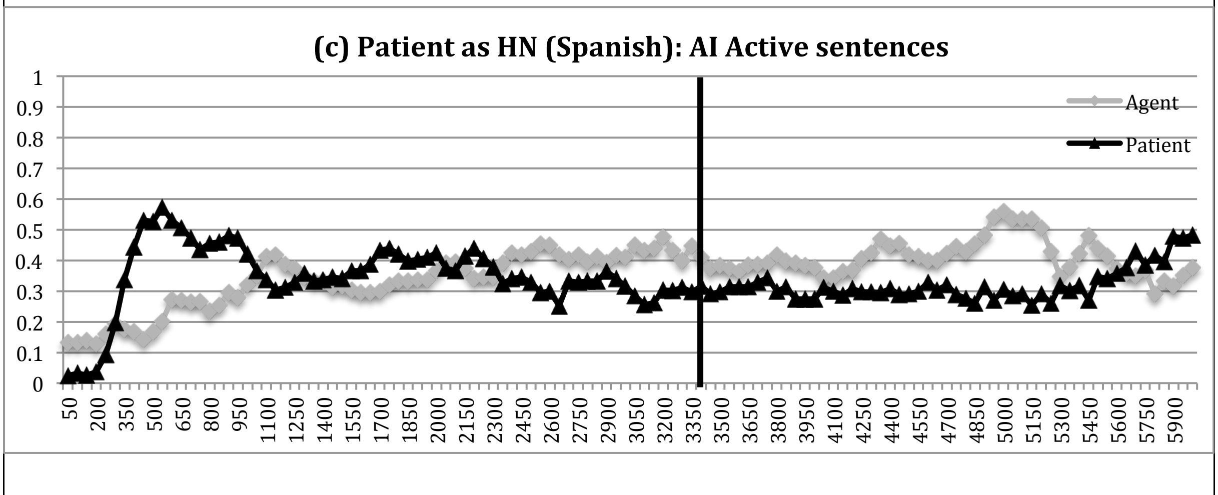


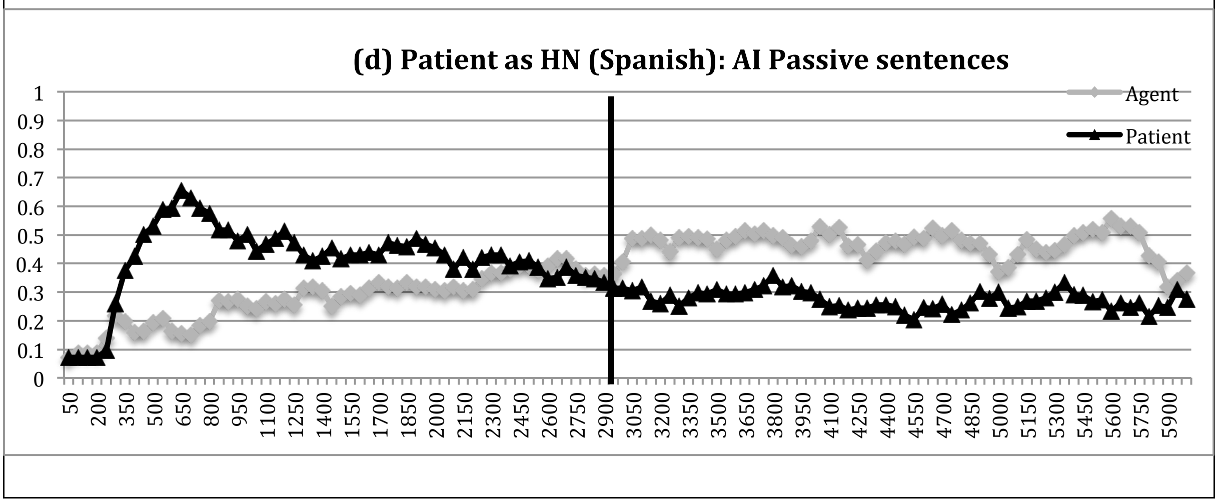
Figure 4.7. Patterns of gazes to agent and patient in Spanish RCs when the patient was the HN, from picture onset until 6000 ms., in all three animacy combinations. Vertical line represents speech onset.

Subsequently, in order to analyze the relation between grammatical assignment and constituent assembly processes (issue number 3 above), we contrasted gaze patterns depending on the voice of the utterance (active vs. passive), focusing again only on sentences with the patient as HN. As explained earlier, these sentences can be uttered in passive or active voice with no changes in word order. Since we did not exert any influence on the voice participants could produce, number of token included in each category varies (see right part –Patient-HN, in Figure 4.4 above, for the different proportions of active and passives). Figure 4.8 below shows the pattern of gazes as a function of voice in the different animacy combinations.




(b) Patient as HN (Spanish): AA Passive sentences
Agent
Patient


(c) Patient as HN (Spanish): AI Active sentences
Agent
Patient


(d) Patient as HN (Spanish): AI Passive sentences
Agent
Patient

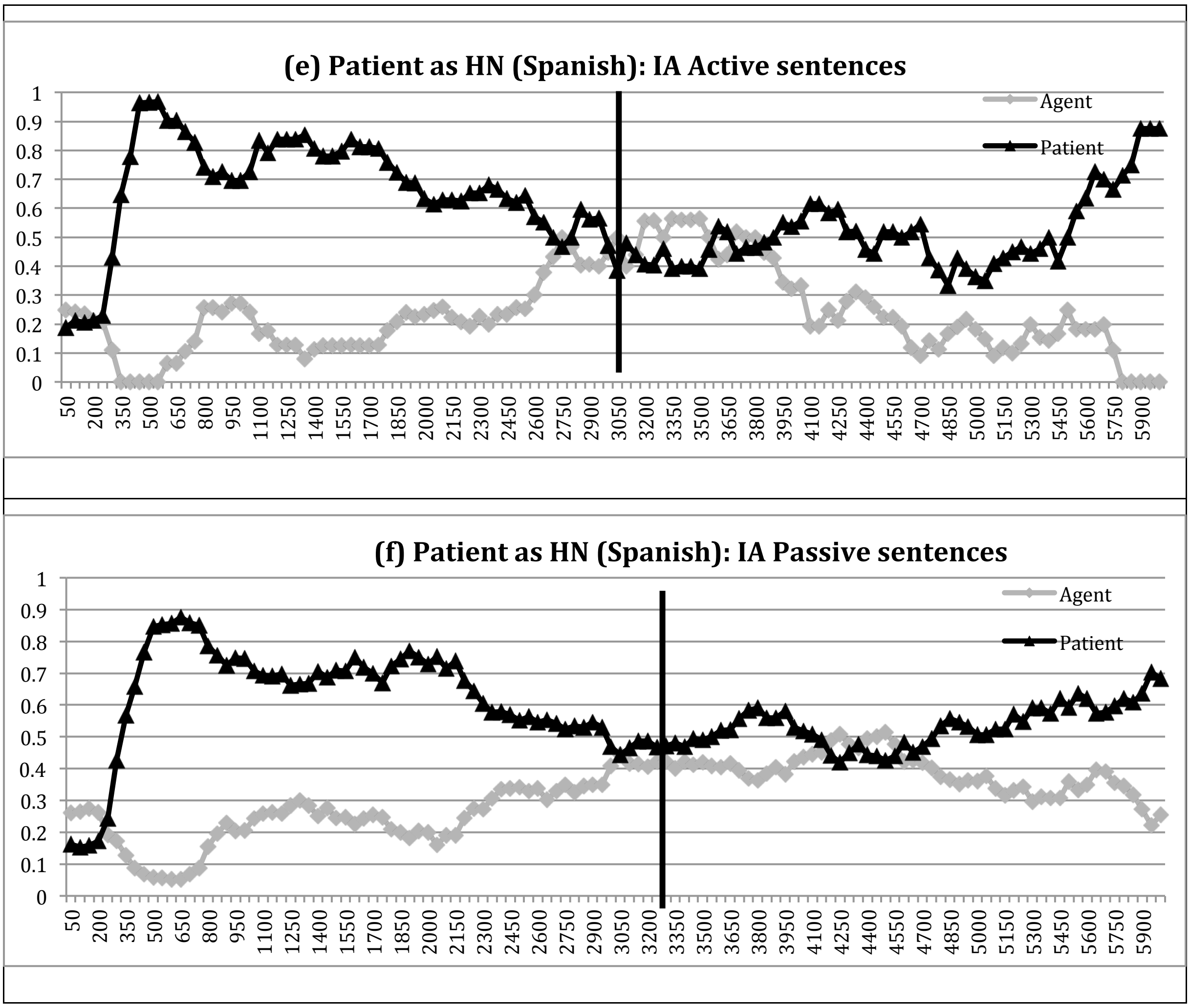


Figure 4.8. Gazes to agent and patient in Spanish RCs from picture onset to 6000 ms. in the Patient-HN condition, in the three animacy combinations, contrasting active and passive responses. Vertical line represents Speech Onset.

The analysis yielded a significant interaction between voice and animacy in TW3 (from 1000 ms. to 1800ms.), in both gazes to agent (t = 3.061, p = 0.002) and to patient (t = -3.19, p = 0.002), and a main effect of voice in patient in TW4 (from 1800 ms. to 2500 ms.) (t = 2.299, p = 0.02), but only marginally significant in gazes to agent (t = -1.737, p = 0.087). In general, there is a convergence of gazes when an active sentence is under preparation, with an increase of gazes to the agent (TW3: active = 0.29, passive = 0.24) and decrease of gazes to the patient (TW3: active = 0.45, passive = 0.53). However, this pattern, and particularly its precise timing, seems to be mediated by the animacy of the items. In TW3, the same pattern of convergence of gazes in active sentences is found for the AA and AI conditions, but the opposite pattern is found in IA sentences (gazes to the agent: active = 0.11, passive = 0.22; gazes to the patient: active = 0.68, passive = 0.55). A closer look at the variation in gaze patterns along time suggests that this is due to an earlier increase in gazes to the agent in IA sentences (Figure 4.7), but these differences start fading in TW4. Taken together, these results suggest that there are differences due to the grammatical function of the elements, even when word order is kept constant.

Discussion

In this experiment, we analyzed the production pattern of RCs in Spanish, both with the agent and the patient as HNs. Results showed a pattern in which speakers tend to look at the element they are going to utter first before its articulation, regardless of its being the object or the subject of the subordinate clause. This general pattern replicates the one found with simple transitive clauses in previous studies (Griffin & Bock, 2000; Konopka & Meyer, 2014). However, there are differences worth noting in the way simple clauses and relative clauses are planned.

First, effects due to RC type are only found before speech onset. In simple transitive clauses, speakers start focusing extensively on the second element at the same time they are articulating the first one. However, in Spanish RCs, although numerically there is a tendency to look at the second element, this tendency is not strong enough, with gazes to the agent and to the patient not differing depending on RC type. The reason for this might lie in the fact that the second element to be uttered is part of a subordinate clause, dependent on the still not prepared (or even omitted) main clause. For example, our participants had to produce sentences like

(5) La chica que empuja al chico lleva una blusa rosa.

The girl who pushed the boy is wearing a pink blouse.

In this case, "boy" is still a subordinate element that depends on the HN "The girl". Similarly, speakers cannot lose sight of the main predicate that says something about the girl: i.e., that she is wearing a pink blouse. Thus, as a result speakers might keep focusing on the HN even after preparation of the lemma is over, in order to accomplish the construction of the main utterance. However, this hypothesis remains to be clarified.

Secondly, another variable we manipulated was the animacy of the agent and the patient of the participants in the event. We did this in order to explore how visual saliency affects the choice of the starting point of the sentence in RC production.

Uttered sentences showed the expected animacy effect reported in the literature on RCs with the patient as HN (Gennari *et al.*, 2012), where events with animate patients yielded a higher rate of passive sentences, regardless of the animacy of the agent. In addition, the analysis of gaze patterns showed a tendency to fixate animate elements in a higher proportion, agents and patients alike. However, the observed tendency suggests that visual saliency is not what guides speakers' gazes initially. Animacy only turned out to be significant at TW2 (from 400 ms.), even despite the fact that speakers start focusing the first element as early as 350 ms. Additionally, we found no relation between the proportion of fixations to the animate element and the likelihood of promoting that element to the subject function (i.e. of producing a passive RC instead of an active), when the patient was the HN. After that, animacy plays a role in interaction with voice, modulating the specific timing of the assignment of grammatical relations. The results reported here for Spanish suggest a speech planning process in which the effects of animacy occur either by virtue of conceptual saliency or due to semantic plausibility, but only show up after apprehension has taken place, and in interaction with grammatical function assignment.

Finally, we explored the timing of grammatical function assignment by comparing sentences with exactly the same word order but different grammatical relations, namely passive sentences (patient / subject (HN) – agent / object) vs. active sentences (patient/ object (HN) – agent / subject). Results showed differences due to the voice of the utterance from 1000 to 2500 ms., although this difference is mediated by animacy. Active RCs are characterized by a pattern in which gazes to the agent increase for a short period, most likely due to the assignment of the subject function to the subordinate NP. These results suggest that the assignment of grammatical functions is not completely set before name-related gazes start. The speech planning process depicted here points to an interactive process, in which grammatical functions are not fully established before linearization processes start. This supports a cascaded incremental planning process.

In order to determine whether RC planning relies more heavily on linear or hierarchical incrementality we need to compare the performance in Spanish with that in Japanese (Issue 2). We will now move on to experiment 2 (Japanese RC production), before turning to compare both languages in the General Discussion.

Experiment 2: Relative Clause production in Japanese

Method

Participants.

Thirty-two Japanese native speakers participated in this study. They were undergraduate or graduate students at Hiroshima University (mean age: 20, range: 18-23). There were 23 women and 9 men. One participant was excluded, since only 6.5 % of his gazes were recorded. All participants reported normal or corrected-to-normal vision.

Materials, Apparatus and Procedure

Same as in Experiment 1. Instructions, questions and the verbs and other lexical items used during the training session were translated into Japanese (e.g. 「誰が黒いドレスを着ていますか」 (Who-NOM black dress-ACC wearing is-Q) "Who is wearing a black dress?" – 「倒す」 "To knock down").

Results

Pattern of spoken responses

Data analysis. Same as in Experiment 1. Excluded responses amounted to 4.57 % of all responses (i.e., a total of 86 responses out of 1917)[41].

Results. Similarly to Experiment 1, items with the agent as HN elicited active Subject RCs in almost all cases, while RCs with the patient as HN gave rise to a wider variety of responses. There was a main effect of RC type and of animacy, as well as an interaction between both factors (all differences $p < 0.01$), with participants producing a higher proportion of passive sentences when the HN was animate than when it was inanimate (with no differences between AA and IA conditions) (Figure 4.9).

[41] In Japanese, contrary to Spanish, starting from the main clause (i.e. "is wearing a read T-shirt"), instead of the RC under analysis, is much less common, when the RC is modifying the subject of the main clause, as is the case in this study. Thus, there were fewer sentences excluded for gaze analysis in Japanese.

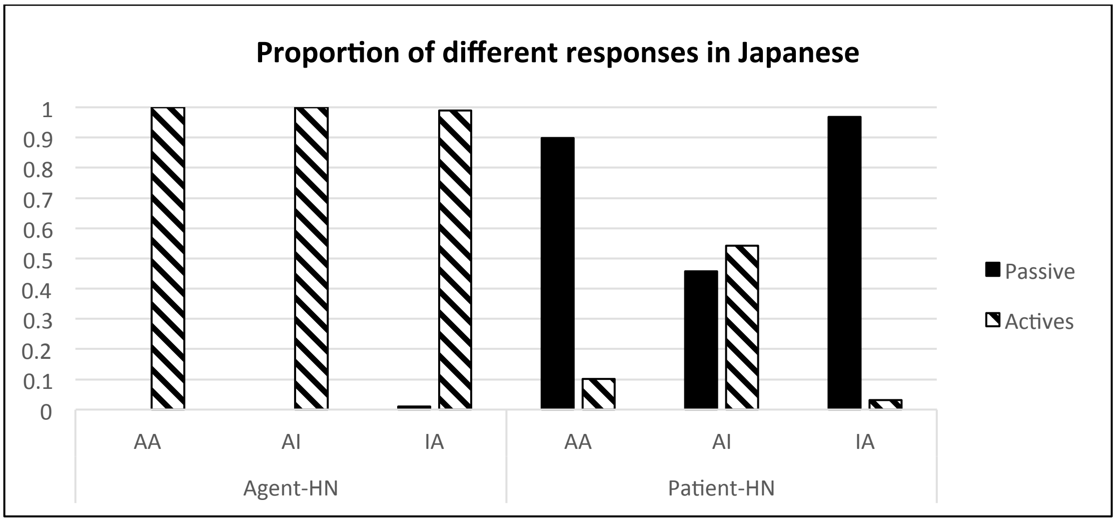


Figure 4.9. Proportion of active and passive responses in Japanese across the different animacy combinations of the Agent-HN condition (left) and the Patient-HN condition (right).

<u>*Pattern of gazes*</u>

Data analysis. The same analysis and Time Windows (TWs) were used as in Experiment 1. In this case, we excluded from gaze analysis the following types of responses:

a. Responses that did not start with the relative clause, mostly sentences with the main clause before the relative clause, thus differing in word order and making difficult any kind of comparison. Examples of this type of sentences will be 「茶色いのは男性を叩いている木です」 – "Chairoi-nowa dansei-o tataiteiru ki desu" – Brown-nominalization-TOP man-ACC is hitting tree is (*What is brown is the tree that is hitting the boy*).

b. Responses in which the HN was uttered before the relative clause. This was the case when participants mentioned first the HN (e.g. "the tree") before reformulating to the whole sentence, with the HN placed at the end. This type of responses was part of disfluencies in speech and was scarce in the data (1.3% of valid sentences). By eliminating this type of sentences we make sure that any early effect or possible fixations in the HN are not due to lexical encoding prior to immediate articulation, since we only accept those sentences in which the subordinate noun is uttered first, then followed by the verb and only with the HN at the end. Responses excluded due to this and the previous criterion together were 25 out of 1755 valid responses (1.42 % of responses).

c. For the same reason as in Experiment 1, we excluded from this analysis responses that failed to overtly mention the two participants involved in the event (e.g. sentences with agent dropped: 叩かれている人 – Tatakareteiru hito (Being hit person) – *The person who is being hit*). Responses excluded due to this condition were 288 out of 1755 valid responses (16.41 % of responses).

d. Finally, we excluded responses in which speech latencies were above 10 seconds or were above or below two standard deviations from each participant's speech onset average.

Under these criteria, excluded responses amounted to 33.28 % of all responses (27.96% of RC responses), with a sample of 31 valid participants, and a total of 1319 responses to analyze.

Results. A general view of gaze proportions along the time-course shows a pattern in which speakers focus on the HN initially, before shifting their gazes to the element that they are going to produce first (the subordinate NP) (Figure 4.10). Speech onset was longer in Japanese than in Spanish, with a mean of 4553 ms. Thus, compared to Spanish, in Japanese there is a delay in starting the utterance that stretches a whole time window (approximately 2000 ms). That said, RCs with the agent as the HN take numerically shorter onset times than RCs with the patient as the HN, although this difference is not significant. There is, however, an interaction between RC type and animacy ($p = 0.022$), due to the fact that sentences with inanimate HNs take longer to begin in both cases: IA sentences when the agent is the HN and AI sentences when the patient is the HN (Table 4.4).

	Agent as HN	Patient as HN			General
		Active	Passive	General	
Animate agent - Animate patient	4363.18	5700.31	4405.60	4552.29	4451.84
Animate agent - Inanimate patient	4339.52	5043.67	4577.41	4863.67	4583.44
Inanimate agent - Animate patient	4752.81	5200.13	4345.21	4377.32	4563.29
General	4473.37	5144	4409.62	4599.32	4533.72

Table 4.4. Speech onset latencies by Japanese monolinguals in sentences with the Agent as HN and the Patient as HN in the three animacy combinations.

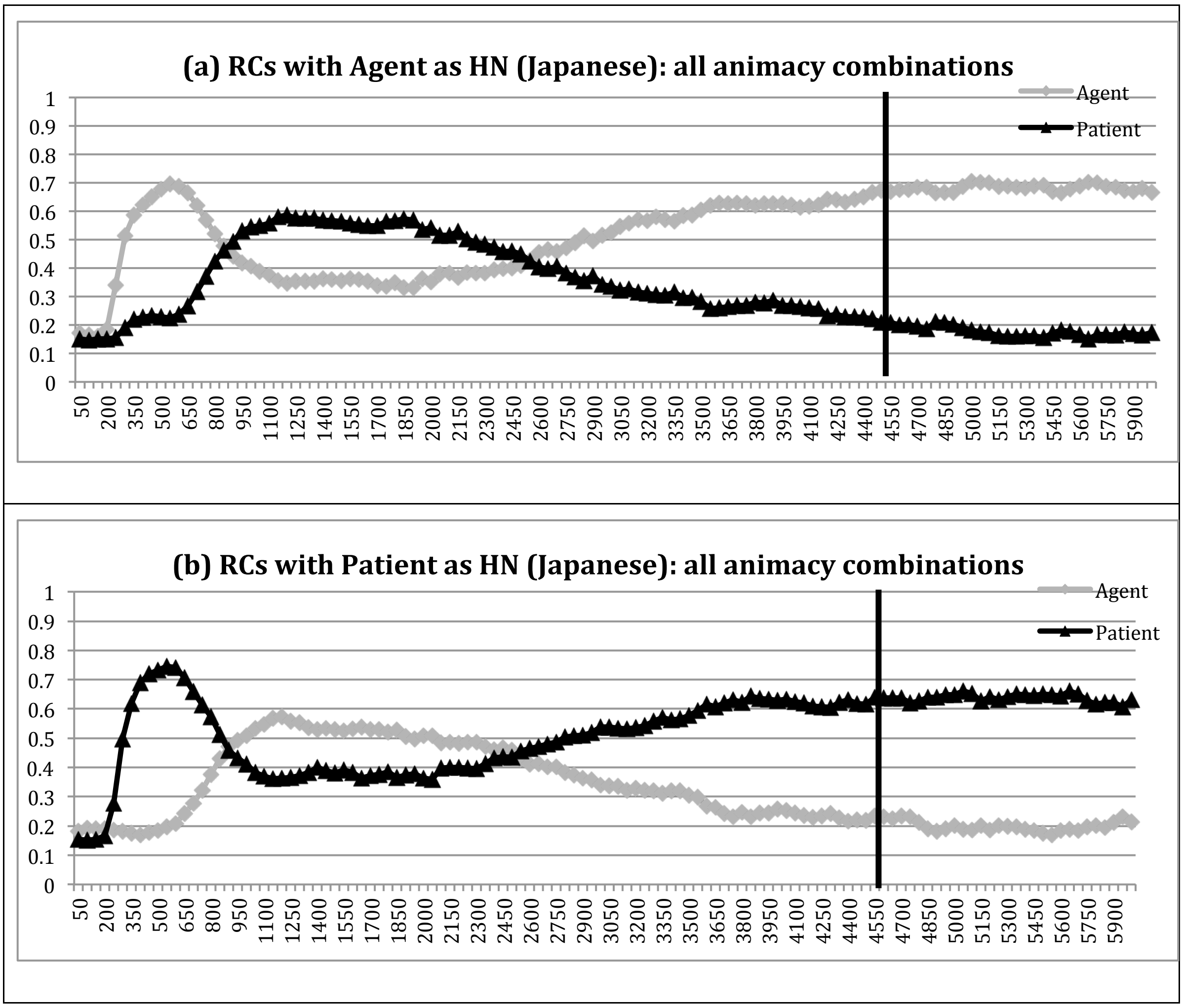
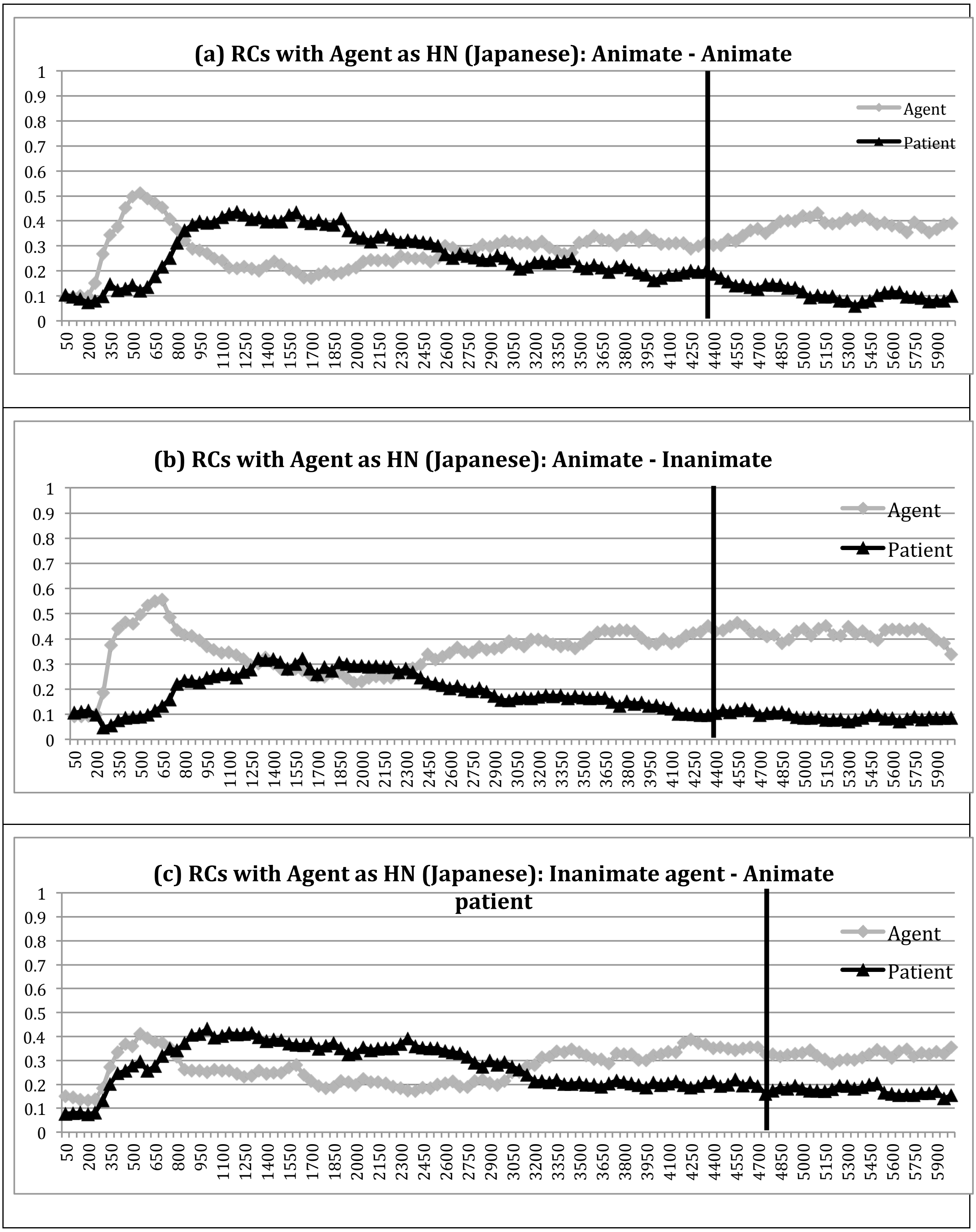


Figure 4.10. General gaze patterns to agent and patient in Japanese RCs with the agent as HN (a) and with the patient as HN (b) (all animacy combinations collapsed): From picture onset until 6000 ms. Vertical line represents speech onset.

Related to issue 1 above, there is an effect of RC type in both gazes to the agent from 0 to 350 ms. (TW1) (t = 4.442, $p < 0.001$) and to the patient (t = -3.611, $p < 0.001$). Participants start to focus their gazes on the element that is going to be the HN from about 300 ms. From that moment onwards, there is a main effect of RC type in all analyzed TWs, with different patterns depending on the TW. Thus, from 400 ms. to 1000 ms. (TW2) (issue 2 above), the HN is fixated to a greater extent, even though it is not placed at initial position (agent: Agent-HN = 0.57, Patient-HN = 0.31, t = 6.787, $p < 0.001$; patient: Agent-HN = 0.36, Patient-HN = 0.60, t = -6.819, $p < 0.001$). From 1000 to 2500 ms. (TWs 3 and 4), participants shift their gazes to the element that is placed at sentence-initial position (i.e., the subordinate NP) (TW3: agent: Agent-HN = 0.36, Patient-HN = 0.54; t = -3.792, $p < 0.001$; patient: Agent-HN = 0.57, Patient-HN = 0.37; t = 3.872, $p < 0.001$; TW4: agent: Agent-HN = 0.39, Patient-HN = 0.49; t = -2.51, $p = 0.016$; patient: Agent-HN = 0.50, Patient-HN = 0.40; t = 2.422, $p = 0.019$). After this brief focus on the first element, participants shift again to the HN, which is focused extensively even after SO (TW5: agent: Agent-HN = 0.58, Patient-HN = 0.31; t = 4.864, $p < 0.001$; patient: Agent-HN = 0.31, Patient-HN = 0.58; t = -4.956, $p < 0.001$; TW6: agent: Agent-HN = 0.68, Patient-HN = 0.21; t = 11.45, $p < 0.001$; patient: Agent-HN = 0.18, Patient-HN = 0.63; t = 10.92, $p < 0.001$). This pattern suggests that participants prepare tentatively the utterance by focusing on the HN before starting lexical retrieval. The tendency to go back to the HN quite early (despite its being the second uttered element) shows a strong reliance on the HN when planning RCs. Interestingly, although numerically there is a tendency to focus on the animate element, regardless of thematic role or grammatical function in the sentence (that is, gazes to the agent: AI > IA > IA and the opposite in gazes to the patient: IA > AA > AI), this tendency is only visible from TW2, and animacy only reaches significance from 4500 to 6000 ms. (TW6) in gazes to the agent (t = -2.048, $p = 0.045$) and only marginally from 2500 to 4500 ms. (TW5) (t = -1.88, $p = 0.06$). Interestingly, there is a marginal main effect of animacy in gazes to the agent in TW1 (t = 1.724, $p = 0.096$), which reaches significance when analyzing only sentences with patient as the HN (t = 2.413, $p = 0.019$). However, the tendency differs here, with a higher proportion of gazes directed at the agent when it was inanimate (IA condition), thus showing difficulties in integrating this element as agent. Moreover, taking into account sentences with the patient as HN, there is no relation in TW1 between the proportion of gazes directed to animate patients and the likelihood of producing a passive sentence (issue 1). In general, there is a tendency to concentrate more gazes in agent and patient *alike* when a passive sentence is going to be produced (significant only for gazes to the patient: t =2.003, $p = 0.049$), but this is not mediated by the animacy of the elements (ts < 2 for interaction between animacy and voice in TW1 in agent and patient) (Figure 4.11).



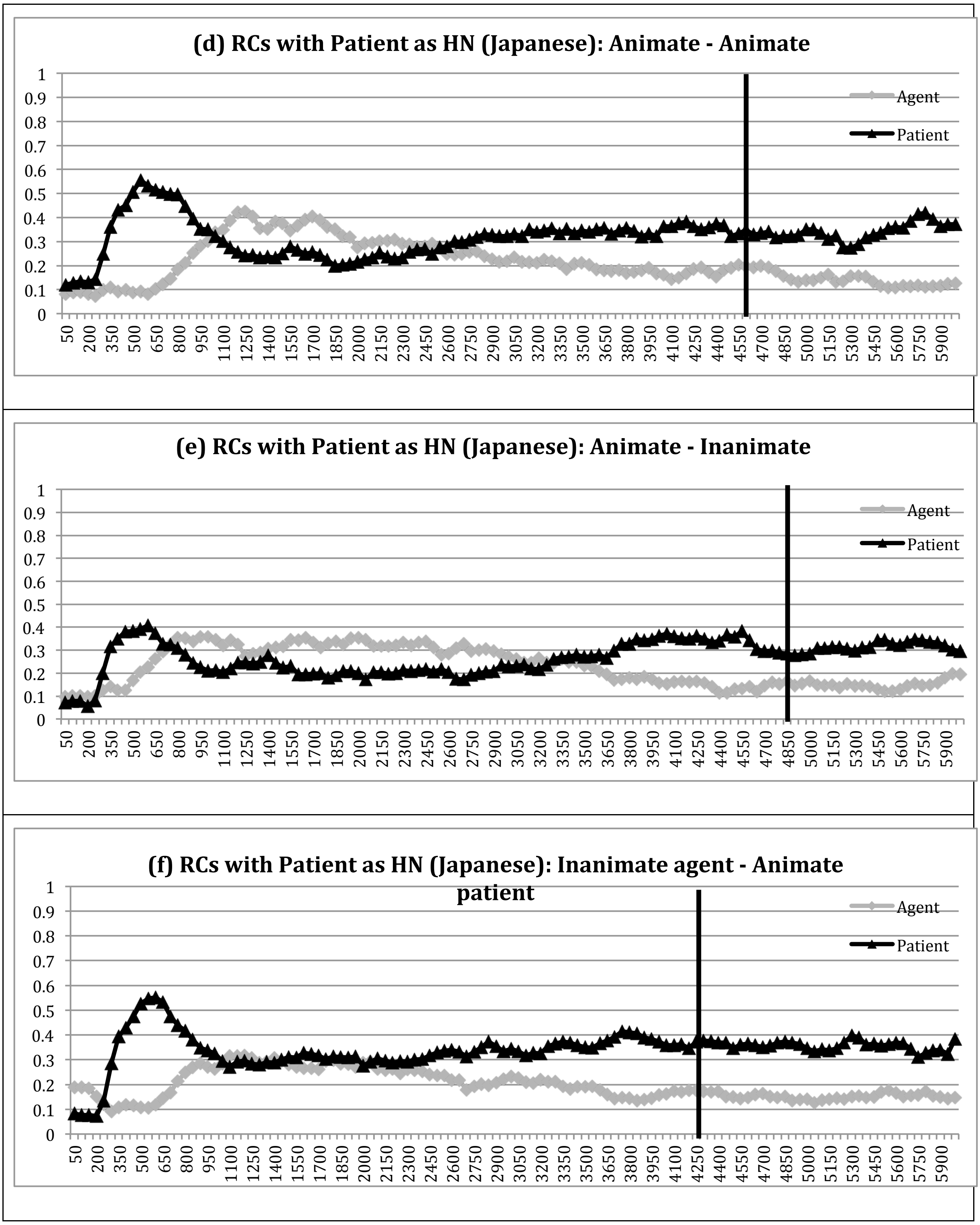


Figure 4.11. Patterns of gazes to agent and patient in Japanese RCs when the agent was the HN ((a) to (c) panels) and when patient was the HN ((d) to (f) panels), from picture onset until 6000 ms., in all three animacy combinations. Vertical line represents speech onset.

Importantly, since speech onset was later in Japanese than in Spanish, there arises the possibility that the pattern found was due to this delay. In order to check whether the pattern found was due to long speech onset latencies, we analyzed the gaze patterns of participants with speech onsets between 1000 and 3000 ms. This time, the SO mean went down to 2328 ms. (agent-HN: 2311 ms.; patient-HN: 2346 ms.). As can be seen in Figure 4.12, the general pattern did not change when taking only into account short speech onset latencies: participants do focus on the HN from 300 to 850 ms. on average, before turning their gazes to the first uttered noun, which is not the one that is focused first.

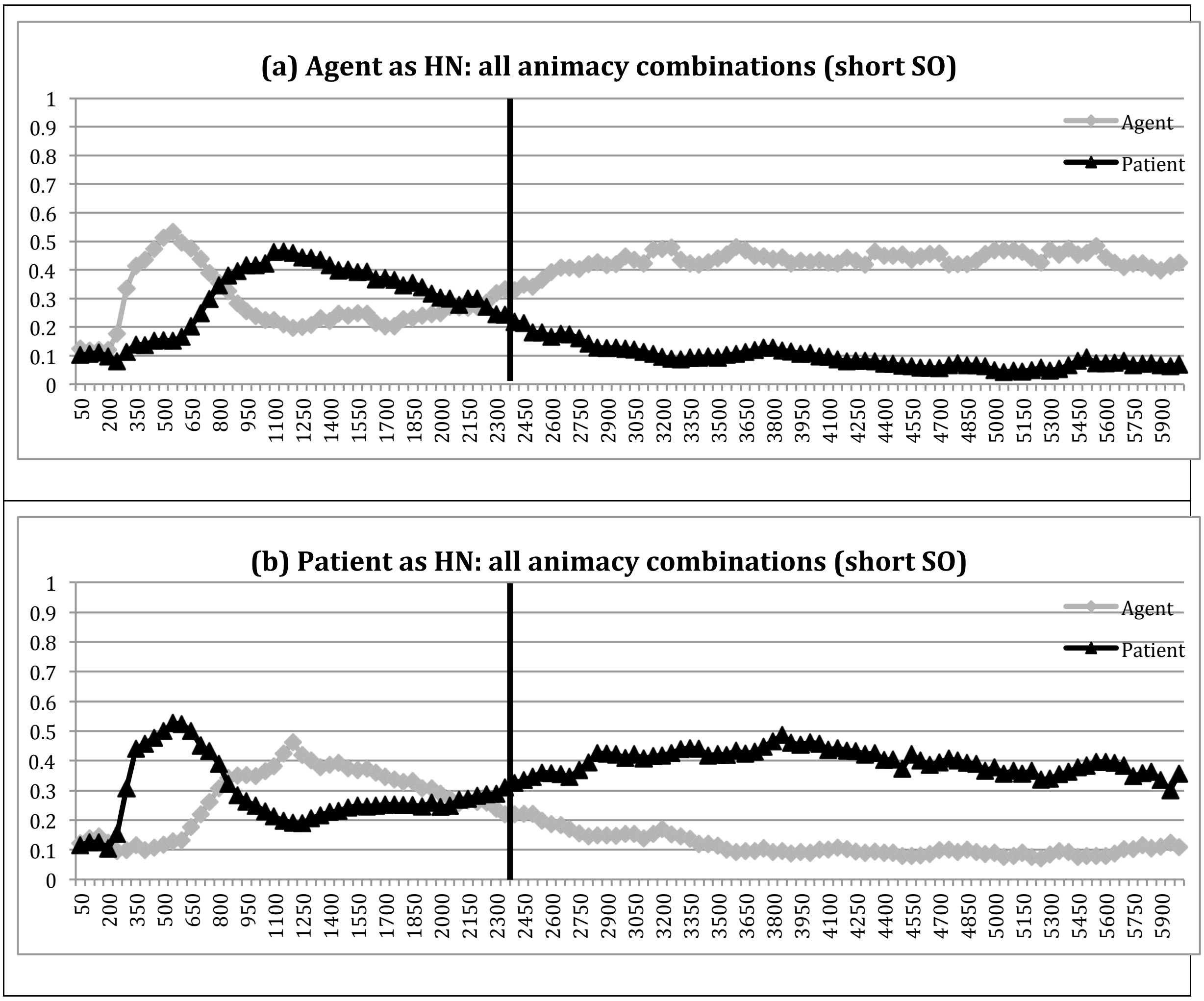
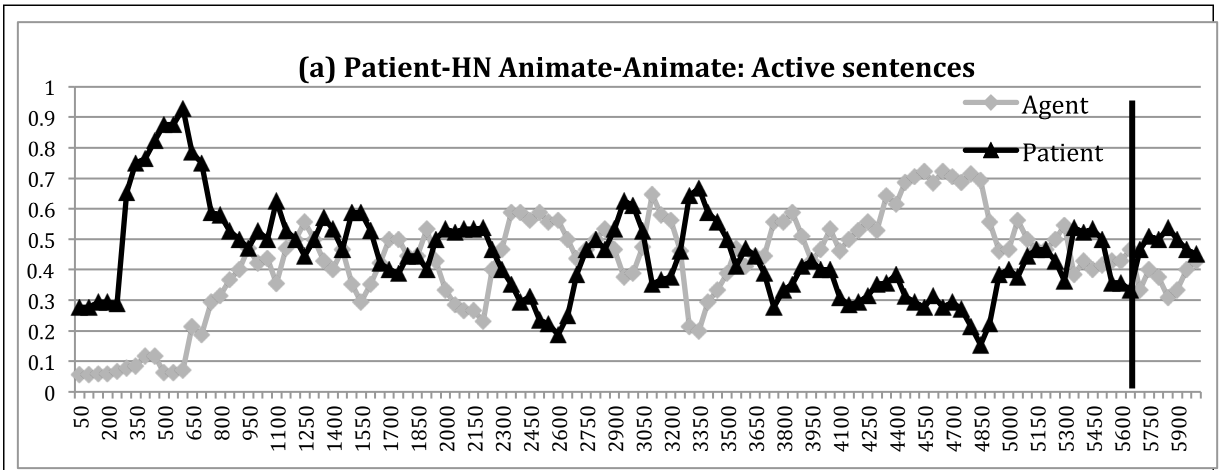
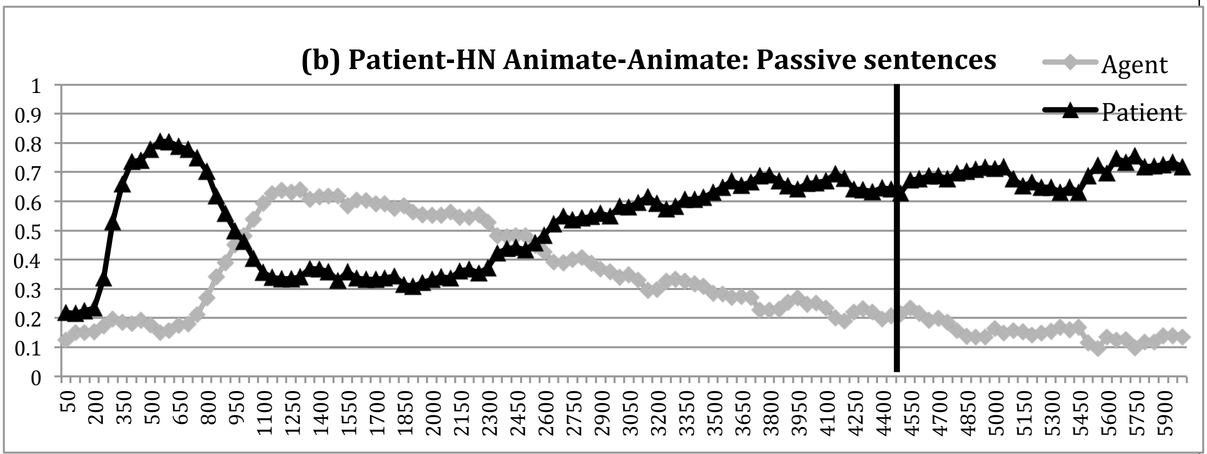


Figure 4.12. Gaze patterns to agent and patient in Japanese RCs with the agent as HN and with the patient as HN (all animacy combinations collapsed), only in responses with speech onsets from 1000 ms to 3000 ms.: From picture onset until 6000 ms. Vertical line represents speech onset.

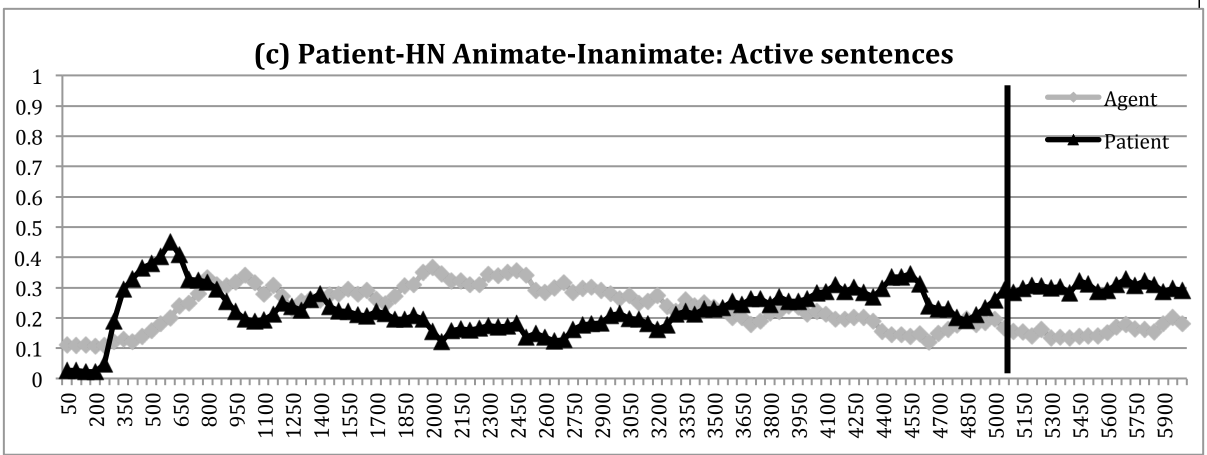
Finally, as we did previously in the Spanish experiment, in order to analyze the effects of voice (active or passive) on speech planning (issue 3), we focused specifically on RCs with the patient as HN (Figure 4.13). Note that, once again, the proportion of responses included in each category varies depending on animacy, as a function of speakers' preferences (see right side –Patient as HN on Figure 4.9 above).



(a) Patient-HN Animate-Animate: Active sentences
Agent
Patient



(b) Patient-HN Animate-Animate: Passive sentences
Agent
Patient



(c) Patient-HN Animate-Inanimate: Active sentences
Agent
Patient

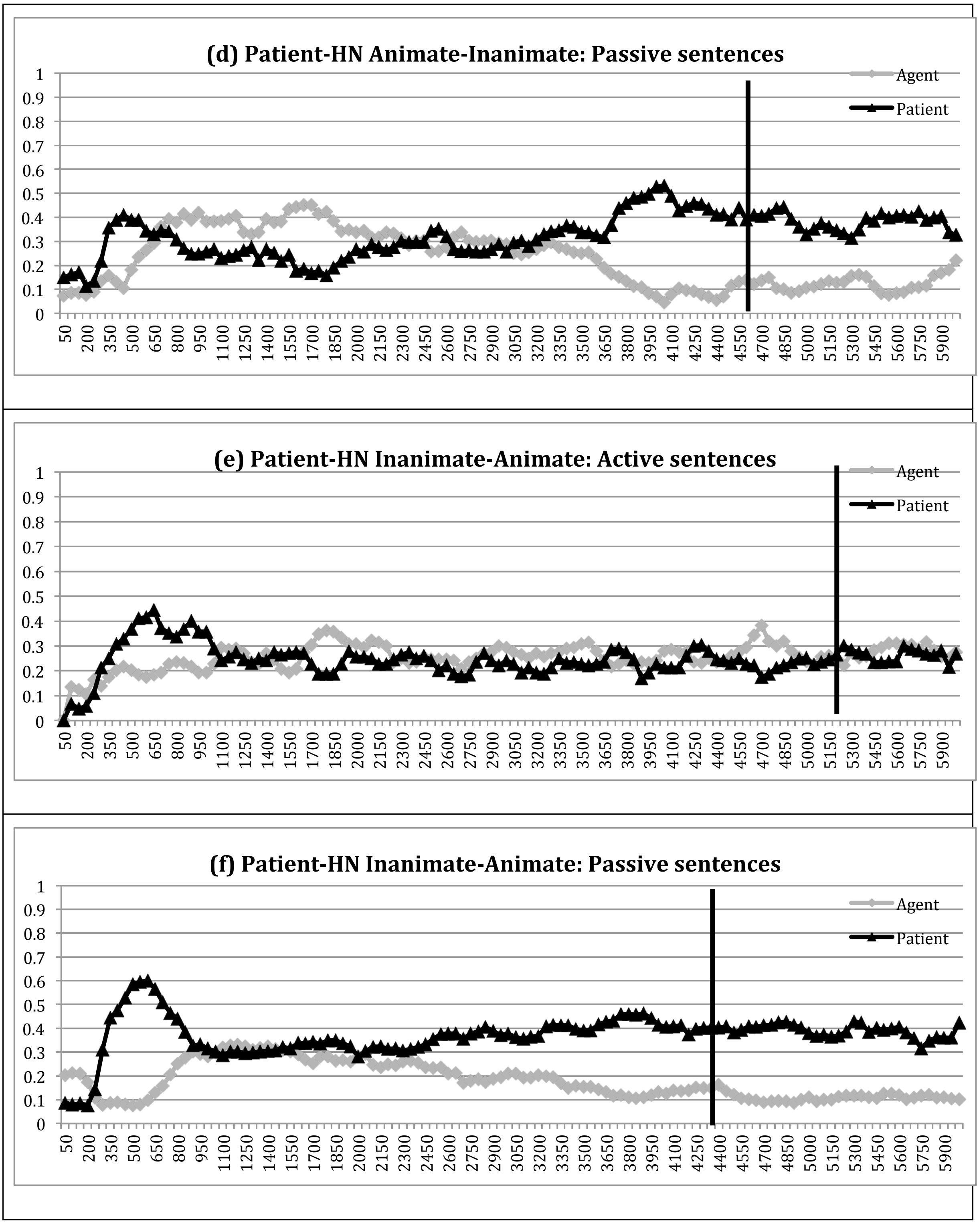


Figure 4.13. Gazes to agent and patient when produced sentences in Japanese were active and passive in the Patient-HN condition, in all animacy combinations: from picture onset to 6000 ms. Vertical line represents Speech Onset.

Apart from the main effect of voice in gazes to the patient in TW1 reported earlier, this analysis showed a significant main effect of voice in TWs 5 and 6 (from 2500 ms.) in both gazes to the agent (TW5: t = -3.734; $p < 0.001$; TW6: t = -4.011; p

< 0.001) and to the patient (TW5: $t = 3.99$; $p < 0.001$; TW6: $t = 3.453$, $p < 0.001$). The pattern found, both prior to and after SO, is that gazes directed to the agent increase when the to-be-uttered sentence is active, with the opposite pattern in gazes to the patient, resulting in a pattern of convergent gazes in active sentences, in comparison to passive sentences (TW5: agent: active = 0.44, passive = 0.27; patient: active = 0.44, passive = 0.62; TW6: agent: active = 0.33, passive =0.16; patient: active = 0.52, passive = 0.66).

Discussion

In Experiment 2, we examined the production of RCs in Japanese, a head-final language. In this language, as explained above, the HN of the RC is placed at the end of the subordinate clause, resulting in a pattern in which the most dominant element of the sentence cannot be uttered until the end of the clause.

Following the same procedure as in Experiment 1, we first analyzed the effect of animacy on RC production. Replicating the results found in Montag & MacDonald (2009), we found an effect of animacy on spoken sentences: in sentences with the patient as HN, there was a tendency to promote animate patients to the subject function, regardless of the animacy of the agent. On the other hand, when analyzing gaze patterns, we found that animacy by itself (i.e., visual prominence) had no discernible influence on gaze patterns to the agent or the patient along the whole period of time monitored. Numerically, there was a tendency to fixate animate elements in general, regardless of their being the agent or patient of the sentence. However, this tendency was only visible from TW2 and was significant only after speech onset, which suggests that the visual saliency of the items was not what guided planning and caused the animacy effects observed on produced sentences. Moreover, in RCs with the patient as HN, we observed no relation between the proportion of fixations to the animate HN and the likelihood of promoting it to subject function, suggesting once again that is was not the animacy of the isolated element what determined the final shape of the utterance. Rather, IA elements attracted more gazes in general during the first ms. after picture onset. This is especially true for inanimate agents, as shown in the results. As we saw in chapters 1 and 3, Japanese is a language that relies heavily on animacy cues, showing low acceptability of inanimate agents (Rosen, 1999). For that reason, IA sentences present a challenging situation in which the whole scene is difficult to map onto linguistic form. This fact points towards a global encoding of the scene in a period as short as 300 ms. Additionally, from TW1 there were differences due to RC type, since participants started fixating the HN as soon as 300 ms. These results suggest that the understanding of the scene takes place in a very short period (Griffin & Bock, 2000).

From 300 to 1000 ms., gazes were directed to the HN, despite its not being the first uttered element. These effects might be interpreted as showing that the structure of the RC is being selected at this early stage, with participants being guided by the plausibility of the whole scene rather than the animacy of single elements. This would explain the lack of animacy effects on gaze patterns, as well as the animacy effects found in uttered sentences (i.e. the tendency to promote animate elements to subjects due to their greater prototypicality). We will come back to the implications of these results in the General Discussion of this dissertation (chapter 6).

Once the structure is prepared (after 1000 ms.), gazes are directed to both items of the picture in order to retrieve their corresponding lexical items in preparation for speech. However, the preparation of the first element takes place quickly with participants shifting again to the HN approximately two seconds before starting to speak. Thus, in Japanese, gaze patterns reveal a process in which the whole utterance seems to be prepared before articulation begins, thereby yielding longer speech onset times. However, a closer look to the increased speech onset average indicates there is a large variability between participants. Japanese speakers, in general, prefer to forego fluency, adopting a strategy that favors response accuracy over incremental production. However, this probably results from a conscious strategy (Wagner *et al*., 2010; Konopka, 2012) rather than from the inherent planning process required by the structure under preparation (i.e. RCs in Japanese). In order to test whether the pattern found for Japanese RC planning was due to a strategy or to processes of incremental planning, we analyzed the pattern of gazes of our faster participants only, by considering only responses in which SO was between 1000 ms. and 3000 ms. We observed that fast speakers did not show the general delay that was observed when taking into account all sentences: they started to speak as soon as they prepared the first element of the sentence, and at the same time that they started retrieving the second one. But importantly, the initial process of focusing the HN, as well as its duration, did not differ in both analyses, thus replicating the main findings of this study. The reasons why a large amount of participants opted for accuracy over fluency might have had relation with cultural characteristics, along with the fact that we did not restrict the time they could spend in each item (we did not restrict it in any of the three studies introduced here), but the initial structural processing remained identical in both groups of fast and slow speakers.

Finally, once again taking into account both fast and slow speakers' responses, we observed voice effects were significant right before SO, showing a late effect of grammatical function assignment in Japanese as well.

We will turn to the comparisons with the Spanish results in the next section.

Comparison between Spanish and Japanese production

In this chapter, we have reported two experiments comparing the planning processes involved in RC production in a head initial language and a head final language from an online perspective. We first analyzed the effects of conceptual saliency (i.e. animacy) on RC production, to further relate it to three specific issues addressed in this study, namely (1) the onset of sentence planning and the role of visual saliency in its initial stages; (2) the information that is prioritized in linguistic encoding, either structural or lexical; and (3) the relationship between the processes of grammatical function assignment and constituent assembly during sentence planning. In the final general discussion, we will try to recap what can be concluded from these two studies regarding these three issues in relation with our hypotheses and aims and relating it with our next study with bilingual speakers. To avoid repetition, here we will briefly introduce what can be concluded from these two experiments focusing especially on those hypotheses that focused on the comparison between languages:

1. *Apprehension stage:* In both languages, the animacy of the elements had an influence on the uttered sentence, with a tendency to promote the animate element to the subject position, regardless of word order (initial in Spanish and final in Japanese), thus replicating previous studies that analyzed the effect of animacy on RC production in Japanese and Spanish (among other languages) (Montag & MacDonald, 2009; Gennari *et al.*, 2012). However, this tendency was not reflected in gaze patterns in either language. In both cases, there was a tendency to focus more extensively on animate elements, but this tendency did not show up from picture onset, nor was it correlated with the selected voice in

RCs with the patient as HN. Quite on the contrary, there was a tendency to fixate more on both the agent and the patient of the IA condition during the first 350 ms., showing greater costs in interpreting and understanding this kind of events in order to undertake linguistic encoding. This pattern is especially visible in Japanese.

During the first milliseconds, speakers of both languages do not focus preferentially on any of the elements, showing a general pattern in which they try to apprehend the whole scene and look for the element that they are being asked about. This search takes as short as 300 to 350 ms., with Japanese speakers being slightly faster than Spanish speakers in beginning to fixate the main element. Differences regarding timing are small and they might just be due to individual differences of the tested groups. We will go back to this issue in the General Discussion, when we compare these results with those of bilingual speakers.

These results point toward an apprehension stage that is not guided by the dominance of isolated elements, but of the global scene, which will guide gazes occurring from 350 ms. onwards. We will develop this idea with more detail in the General Discussion.

2. *Grammatical encoding:* The main comparison between languages was carried out at this stage. As was explained in the hypotheses section (see Table 4.1), different accounts of incrementality are thought to hypothesize different outcomes. Hierarchical incrementality accounts predicted initial planning to be identical between Spanish and Japanese, despite word order. On the other hand, linear incrementality accounts predicted that both languages would differ from the beginning, since linguistic planning takes place along with the retrieval of the lexical elements in order. Results showed a planning process that was identical between both languages until 1000 ms. This resemblance supports a hierarchical planning account, with participants creating the structure of the relative clause before lexical retrieval in order starts. However, it is important to note that this does not mean that in Spanish there is no lexical retrieval involved in the period from 300 ms. to 1000 ms., since in this language the first mentioned element and the HN match. Thus, in Spanish there is the possibility that planning takes place in a pattern somewhat closer

to linear incrementality, with a more interwoven planning of both relational and non-relational information. We will go back to the implications and limitations of these results in the General Discussion.

3. *Planning during lexical access:* Our results suggest that there is an early encoding of the structure on hand, but that does not entail that grammatical roles are fully assigned at this point, as the differences due to voice (i.e. to the assignment of grammatical role) seem to suggest. In Table 4.2 above, we hypothesized that if grammatical roles were fully assigned at the beginning, or if they were not differentiated from the retrieval of the lexical elements, active RCs should not differ from passive RCs (in both cases, with the patient as HN) in Japanese and in Spanish alike. This would have been the result of actives and passives RCs word order not differing within languages. However, our results showed that, despite having exactly the same word order, there are differences due to the different grammatical role assigned to the agent or the patient. Interestingly, these differences were more marked in Japanese, with effects of voice lasting even after SO. In Spanish, however, effects are restricted to the TW right before SO. The tendency in general is to reduce the differences in gaze proportion between agent and patient in active sentences, might be due either to the assignment of the subject role to agents, or to an increased complexity in the assignment of grammatical role in active RCs depicting non-prototypical (IA) events or events with potentially competing elements (AA). Once again, we will come back in more detail to these results in the General Discussion.

In the next chapter, we will present a study run with Spanish (L1)-Japanese (L2) bilinguals in the production of RCs. In that study, we will try to explore how bilingual speakers deal with different planning strategies when speaking in their L2. A global appraisal of the results of the experiments included in this dissertation will follow in Chapter 6.

Chapter 5. Study 3: Relative Clause production in Japanese by Spanish-Japanese bilinguals

In the present study, we explore how sentence formulation is carried out by bilingual speakers using their L2. In particular, we analyze the time course of relative clause planning in Japanese (L2)[42] by native speakers of Spanish. As reviewed before, word order is entirely different in these two languages, which will allow to explore how bilingual speakers adapt their planning strategies to the demands of each language.

If we take into account the results presented in the previous chapter, we may hypothesize that Spanish-Japanese bilinguals will need to adapt to the head final condition of their L2 by changing their planning strategies. This re-adaptation is, nevertheless, confronted by the narrowing of the planning scope they suffer due to greater use of cognitive resources, as has been reported in monolingual speakers facing difficult cognitive tasks (Wagner, *et al*., 2010; van de Velde & Meyer, 2014; Klaus, Mädebach, Oppermann & Jescheniak, 2017).

On the other hand, when facing the task of planning an utterance in an L2 not yet in full command, bilinguals rely to a greater extent on semantic cues. Prototypicality in the animacy of thematic roles (i.e. animate items being assigned more often an agentive role and inanimate items a patient role) is used in order to reduce the cognitive load demanded by structural planning processes. These strategies of reliance on animacy may help bilinguals overcome the difficulties coming from the advanced planning of the HN.

Thus, in this study we explore how highly proficient Spanish-Japanese bilinguals plan RCs in their second language. We focus on the features that characterize bilingual speakers' speech planning in comparison with monolingual speakers, both in the use of semantic cues (animacy) and in the extent to which they rely either on a structurally-driven planning, or on a lexically-driven planning. We analyze how changes in planning strategies due to the requirements of the output language (Spanish vs. Japanese) are accommodated, and if these shifts entail an additional burden on the planning process of bilingual speakers.

[42] Note that for all these participants, Japanese is not the L2 but the L3 (following English) and in some cases the L4. We will use, nonetheless, the term L2 to refer to any language acquired later than the native language and with less dominance than the first one.

Based on the results presented in the previous chapter with monolingual speakers of Spanish and Japanese, and on previous findings with bilingual speakers, we present our hypotheses as follows:

1. ***Behavioral responses:***

Similarly to what we found in Rodrigo (2013) with bilinguals, and what has been reported in previous studies with monolinguals (Gennari *et al*., 2012; Montag & MacDonald, 2009), we expect that bilingual speakers will make use of animacy cues to facilitate sentence planning. We expect that sentences with prototypical animacy combinations (i.e. animate agent and inanimate patient) will have fewer overall errors and shorter speech onset latencies than less prototypical animacy combinations (i.e. inanimate agent and animate patient) or more potentially ambiguous combination (i.e. animate agent and patient), in accordance with the Competition model (MacWhinney, 1997) which posits that bilinguals will make use of animacy cues so as to ease the task of planning in their L2. At the same time, errors will likely reflect a tendency to assign the animate item the nominative case marker and the inanimate item the accusative case, replicating previous findings.

2. ***Gaze patterns:***

2a. Apprehension stage: In the two previous studies with monolinguals, we saw that in both Japanese and Spanish, speakers showed a very short initial period presumably devoted to understanding the scene they were looking at and identifying the item that would become the head noun of the relative clause. At this stage, the animacy of the individual elements did not exert any influence on the final shape of the utterance. Therefore, we expect to find the same initial stage, which roughly corresponds to the assembling of a conceptual representation of the message to be conveyed, along with the focus of the message at the discourse level (induced by the question given that might tend to highlight the HN). In the case of bilinguals, however, animacy has been reported to play an important role in comprehending and producing relative clauses (e.g. MacWhinney, 1997; Kanno, 2007; Ozeki & Shirai, 2007). Hence, it is reasonable to expect a stronger influence of the animacy of the items on gaze patterns at this initial period than that shown by monolingual speakers. Subsequently, we expect to find no differences due to RC type within the first 300-400 ms.. Nonetheless, the HN will start to be focused within this period, as the speaker grasps the gist of the scene. If animacy has an influence on the initial mapping done by bilingual speakers, we would expect to find differences due to animacy, with participants focusing to a larger extent on animate items and/or focusing on the HN earlier if it is animate in comparison with inanimate HNs. In addition, we anticipate a correlation between gazes to the animate element and a tendency to promote it as the subject in RCs with patient as the HN, either by constructing a passive RC or by erroneously assigning the nominative particle to the object and the accusative to the subject (i.e. reversing the correct grammatical roles to fit animacy prototypicality).

2. b. *Grammatical encoding:* In the two studies presented in the previous chapter with monolingual speakers, we observed a remarkable similarity between Spanish and Japanese speakers from the stage at which they apprehend the scene to the stage at which they start accessing the lexical elements in the same order as they are going to be uttered. In both Spanish and Japanese, speakers shifted their gazes to the HN right after apprehension of the scene. However, for Spanish, this element was the first uttered element, while for Japanese this was not the case. After apprehension, these speakers looked extensively at the item that was the HN of the structure, exactly like Spanish speakers, despite its not being the first uttered element in Japanese. Only after that period, they shifted to the first element to be uttered in order to retrieve the corresponding lemmas to produce the sentence. This suggests that, at least in Japanese, a structurally guided incremental planning occurs, in which a tentative scaffold of the RC is likely created beforehand. In Spanish, however, there arises the possibility that speakers are undertaking structural and lexical planning in a more interwoven fashion, perhaps because their grammar allows so, thus showing a planning process that may be seen as closer to linear incrementality. This difference can have repercussions on Spanish-Japanese bilinguals' incremental production processes.

In the case of Japanese, RCs seem to require a wider planning scope than in Spanish. Relative clauses in head initial languages also can be part of a single scope under certain circumstances (e.g. Lee *et al*., 2013; Ferreira, 1991; Smith & Wheeldon, 1999). However, when cognitive load increases, planning scope narrows, thus shifting scope from clausal to phrasal planning (e.g. Wagner *et al.,* 2010). However, narrowing the scope when producing Japanese RCs would likely result in difficulties in understanding the structure of the output for bilingual speakers (since the main element, the HN, does not show up until the end of the clause). Thus, bilingual speakers will experience a conflict in which the structure of the L2 message they want to convey requires a wider scope, while the increased cognitive load due to L2 planning induces them to narrow their planning chunks. Of course, it might also happen that high cognitive load makes incremental planning impossible for these speakers, forcing them to prepare the whole utterance in advance (e.g. Konopka, 2012; Wagner, *et al*., 2010).

In either case, planning will result in larger speech onsets than those of Spanish or Japanese monolingual speakers. The underlying planning process causing this larger speech onset will become clearer by analyzing gaze patterns. If bilingual speakers opt for the choice of planning the whole utterance before starting production, we will observe that speakers will focus extensively on both agent and patient in the same speech order before starting to speak. If, on the other hand, fluency is favored over accuracy, they might start speaking before planning the whole sentence in advance. If such were the case, and there are still increased speech onset latencies, we could expect differences in the timing of structure building processes, as bilingual speakers adapt to the strategy required by the L2. The nature of these differences is difficult to pin down beforehand, but we could expect both longer time devoted to fixating the HN (not the first uttered item), before preparing the lemmas in the order of speech. On the other hand, we might also expect participants to have problems between deciding to use a lexical or a structural strategy, which will result in gaze patterns in which both the HN and the first uttered element are highly and equally fixated in early time windows. Thus, we will analyze the strategy that bilinguals choose, and its implications on bilingual sentence production.

2c. *Planning during lexical access:* Finally, in studies 1 and 2 with monolingual speakers, we observed that after lexical access started there was still a difference in gaze patterns depending on the grammatical function of the elements involved, which eventually shows up in the choice of sentence voice: in sentences with the patient as HN, the proportion of gazes to the agent was higher when it was located at subject position in active sentences than when it was assigned to the by-phrase in passives. These results occurred both in Spanish and Japanese, although the tendency was stronger in Japanese. This was interpreted as suggesting that early structural planning of the utterance is not complete, but the specific grammatical functions are assigned alongside lemma selection. The stronger tendency found in Japanese might also indicate more difficulty in creating active RCs with less prototypical animacy combinations, due to lower overall frequency of active RCs with the patient as HN, and to greater reliance on conceptual prototypicality. In the case of bilingual speakers, we could expect to find one of the following two scenarios:

a. If bilingual speakers create a structure before lexical retrieval starts, in the same way as Japanese monolingual speakers do, we could expect to find similar effects of voice as in Japanese monolinguals, with differences due to voice showing up from the time window prior to speech onset.

b. On the other hand, if bilingual speakers present difficulties in lexical access, this could lead to a pattern in which gazes are directed to both elements alike during lexical access. If such were the case, we might expect a pattern of gazes that differs from either group of monolinguals, with a less than efficient planning process. If bilingual speakers experience difficulty in lexical access, however, we should not find differences between active and passive voices.

Experiment 3: Relative Clause production in L2 by Spanish-Japanese bilinguals

Method

Participants

Data from 14 participants was recorded at Waseda University (Tokyo, Japan). However, due to low proficiency of some of the participants, as revealed by their score in the SPOT test, only data from 9 was taken into account (those scoring above 70% correct responses).

Bilinguals were all living in Japan and had all been working or studying in Tokyo for a period longer than one year, and used Japanese in their daily life. The mean age of those included in the analysis was 24.33 (range: 21-36). The sample consisted of 6 women and 3 men. One out of the nine participants of this sample was from Venezuela, while the rest were from different regions of Spain. Mean score in the SPOT test for selected participants was 90.26 (range: 100 – 76.92).

Materials

Same materials were used as those in Experiments 1 and 2 with monolinguals speakers. Written instructions, questions and verbs given during the task were presented in Japanese. However, the amount of Kanji characters employed was reduced significantly to accommodate Spanish speakers' knowledge, and to avoid reading comprehension problems to be a factor that could hamper performance. Spoken instructions were given in Spanish and all the interactions with the experimenter (a Spanish native speaker) was held in Spanish.

Apparatus

Same as in experiments 1 and 2.

Procedure

The procedure was very similar to that in Experiments 1 and 2. Before the task, bilingual speakers performed a placement test: the SPOT test for Japanese (Kobayashi, Niwa & Yamamoto, 1996), which consisted of a 7-minute audio with Japanese speech at normal rate (the test is presented in Appendix 5). Participants were asked to fill in the blanks in a text with the information they were listening to in the audio.

After this test, and before the experiment began, participants examined separately all pictures with the animate and inanimate characters/objects and the main actions that would be depicted during the task. They were asked to name them, and when they did not come up with a response, the experimenter provided the correct answer. Since many of the words were new for most of the participants, this examination and rehearsal was repeated several times. The experimenter, nonetheless, encouraged participants to use the noun that was more familiar to them. This was made to make sure participants understood all the items involved in the scenes, and knew a proper way to name them all.

After that, the experiment began and was conducted exactly like Experiments 1 and 2. After the instructions, there were four practice items. The experimenter helped participants to understand the task without providing any correct or model answer. Similar to previous experiments, there was no time pressure, but participants were encouraged to answer with the first and most natural response that came to mind, and to not take long in thinking their reply. They were told that there were no correct responses, so they should not feel the pressure of an exam.

The task was divided in two blocks. Between the first and the second blocks there was a short break.

Results

Pattern of spoken responses

Data Analysis. Spoken answers were recorded and transcribed for analysis. Answers that failed to use a relative clause or whose word order was changed were excluded from analysis. The remaining sentences were classified as valid. Valid sentences were subsequently classified as correct (grammatically correct sentences, even with lexical errors) and incorrect (those with grammatical errors). Incorrect sentences were subclassified as: (1) errors of particle change, or (2) errors of overnominalization (i.e. adding a "no" particle between the subordinate verb and the HN). Errors of particle change were analyzed qualitatively, taking into account the type of particle change and the type of relative clause and animacy combination where it occurred (see description and examples below).

Responses in which participants changed the verb but kept the overall meaning and the transitive structure were classified as correct. Additionally, sentences in which the subordinate NP was omitted were included in the analysis.

A LME model was run for proportion of errors with type of RC (agent as HN and patient as HN) and animacy combination (animate agent-animate patient, animate agent-inanimate patient and inanimate agent-animate patient) as fixed factors. Random factors were kept to their minimum level when no improvement of the model was registered with the inclusion of participant or item variability.

Results. First, we will report the pattern of responses of Spanish-Japanese bilinguals in Japan. As can be seen in the blue bars in Figure 5.1, this group showed a high proportion of valid answers, with an overall 90% of responses suitable for analysis. The only exception is sentences with an inanimate agent as HN. These sentences showed greater difficulty for participants, with only 78% valid responses.

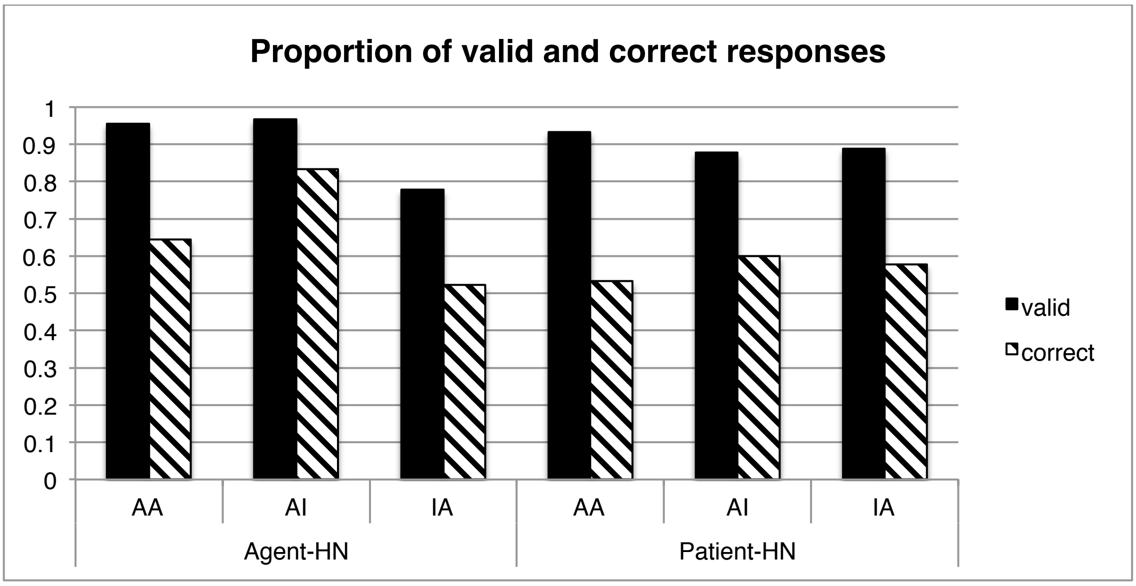


Figure 5.1. Proportion of valid responses (sentences that were uttered as relative clauses with correct word order) and correct responses out of total responses (lexical errors were not taken into account).

Figure 5.1 also shows in the red bars the proportion of correct responses out of total utterances. Overall correct responses amounted to 61.8%, raising to 68.7% in the case of valid responses (i.e. responses that were uttered as an RC). As can be seen, in general terms, sentences with the agent as HN were responded to more accurately, with a higher proportion of correct responses than sentences with the patient as HN (agent-HN: 0.67 vs. patient-HN: 0.57). The highest proportion of correct responses was found in sentences with AI animacy combinations, showing certain awareness or use of the animacy cues when producing relative clauses in Japanese. However, none of these differences was significant using a Linear Mixed Model analysis.

Analysis of the voice of responses in correct responses only showed a high proportion of passive sentences when the patient was the HN (when the agent was the HN, all correct sentences were given in active voice). Similarly to native speakers, there was an effect of animacy: animate HNs resulted in an increased number of passive sentences, despite their clause final position (Figure 5.2).

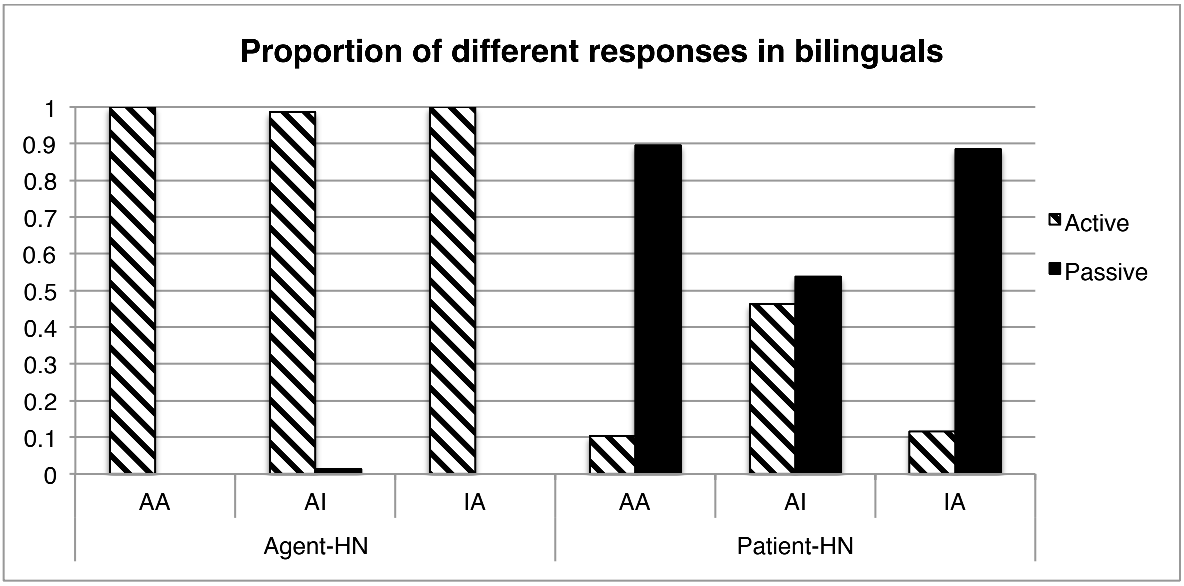


Figure 5.2. Proportion of active and passive sentences out of the complete set of correct responses in the Agent-HN and the Patient-HN conditions.

Turning now to the analysis of the type of errors, we focus on the analysis of errors consisting of changes of case particles in the internal noun phrases (NP). Remember that the so-called internal NP is the first uttered constituent in Japanese RCs. When the agent is the HN, the RC is regularly rendered in the active voice. In such cases, the agent is in clause-final position with the subject role, whilst the patient appears at clause-initial (internal) position with the object role and the accusative particle (を wo). On the other hand, when the patient is the HN, the sentence may be either active or passive: the active form begins with the subject-NP marked with the nominative case particle (が ga); the passive form, in turn, begins with the internal-NP (in this case, the HN) marked with the dative case particle (に ni) (the Japanese equivalent of the by-phrase). Based on the previous description, the following types of errors could be observed:

For sentences with the patient as HN, in active form:

1. Use of the accusative case particle instead of the nominative. For example:

   (1) ええ…釣りのものを引っ掛けているのおじいさんです。(Participant 4)

   Eeh... Tsuri-no mono-o hikkaketeiru-no ojiisan desu.

Eeh... The thing for fishing-ACC is hooking GEN old man is.[43]

(Intended: Fishing rod-NOM is hooking old man is. – "The old man who the fishing rod is hooking").

(2) 農家を押していることは緑カートです。(Participant 10)

Nouka-o oshiteiru koto-wa midori kaato desu.

Farmer-ACC is pushing thing-TOP green cart is.

(Intended: Farmer-NOM is pushing thing-TOP green cart is. – "The thing that the farmer is pushing is a green cart").

2. Use of the dative instead of the nominative.

(3) バスに追いかけている警察が…シャツが黄色いです。(Participant 3)

Basu-ni oikaketeiru keisatsu-ga... shatsu-ga kiiroi desu.

Bus-DAT is chasing police-NOM... shirt-NOM yellow is.

(Intended: *Active*: Bus-NOM is chasing policeman-TOP shirt-NOM yellow is. – "The policeman who the bus is chasing wears a yellow shirt" or *Passive*: Bus-DAT is being chased policeman-TOP shirt-NOM yellow is. – "The policeman who is being chased by the bus wears a yellow shirt").

(4) ええと…警察にとらえしているの泥棒がオレンジの T シャツを着ています。(Participant 2)

Eeto... Keisatsu-ni torashiteiru-no dorobou-ga orenji-no T-shatsu-o kiiteimasu.

Eeh... Police-DAT is catching-GEN thief-NOM orange-GEN T-shirt-ACC is wearing.

(Intended: *Active:* Policeman-NOM is catching thief-TOP orange-GEN T-shirt-ACC is wearing. – "The thief who the policeman is catching is wearing an orange T-shirt". *Passive:* Policeman-DAT is being caught

[43] The inclusion of the genitive particle between the subordinate verb and the HN is also a very common error between participants, which was also present in learners in Rodrigo (2013) study.

thief-TOP orange T-shirt-ACC is wearing. – “The thief who is being caught by the policeman is wearing an orange T-shirt”).

3. Use of the postpositional adverb “kara” (from) instead of the nominative.

   (5) ええ…警察官から捉えている泥棒はオレンジシャツを着ています。(Participant 5)

   Ee... Keisatsukan-kara toraeteiru dorobou-TOP orenji shatsu-o kiiteimasu.

   Eeh... Policeman-FROM is catching thief-TOP orange shirt-ACC is wearing.

   (Intended: *Active:* Policeman-NOM is catching thief-TOP orange T-shirt-ACC is wearing. – “The thief who the policeman is catching is wearing an orange T-shirt”. *Passive:* Policeman-DAT is being caught thief-TOP orange T-shirt-ACC is wearing. – “The thief who is being caught by the policeman is wearing an orange T-shirt”)[44].

4. Use of the particle “de”, which marks instrument role (similar in meaning to “with”) instead of the nominative.

   (6) トラックで運べているの男の人は緑の制服を着ています。(Participant 2)

   Torakku-de hakobeteiru-no otoko-no-hito-wa midori-no seifuku-o kiiteimasu.

   Truck-INST is carrying-GEN man-TOP green-GEN uniform-ACC is wearing.

   (Intended: Truck-NOM is carrying man-TOP green-GEN uniform-ACC is wearing. – “The man who the truck is carrying is wearing a green uniform”).

On the other hand, when these sentences were uttered in passive form, the errors regarding particle change slightly differ. The following was observed:

---

44 We assume that the use of kara, which means “from” is an interference from the L1. In Spanish passive sentences, the preposition “by” used in passive sentences is translated as “por”, which also translates as “from” or “kara” in Japanese. Thus, we assume that our participants could have been trying to use a passive. However, since the subordinate verb is in active form, the possibility that they changed only the particle or that they mixed both possibilities also stands as plausible. This also applies to errors of using dative while keeping verbs in active form, as happens in the error just prior to this.

1. Use of the accusative instead of the dative.

    (7) 警察を照らせている物は丸です。(Participant 4)

    Keisatsu-o terasareteiru mono-wa maru desu.

    Police-ACC is lighting up thing-TOP rounded is.

    (Intended: Policeman-DAT is lighting up thing-TOP rounded is. – "The thing that is being lighted up by the policeman is rounded").

2. Use of the postpositional adverb "kara" (from) instead of the dative.

    (8) ええと…ええ…んん…警察から照られ…照らせた物は丸いです。(Participant 2)

    Eeto...Ee...Nn.. Keisatsu kara terasare... terasareta mono-wa marui desu.

    Eeeh... From the police is being lighted up thing-TOP rounded is.

    (Intended: Policeman-DAT is being lighted up thing-TOP rounded is. – "The thing that is being lighted up by the policeman is rounded").

In sentences with the agent as HN, errors were less varied, with only the following kinds:

1. Use of the passive form instead of active verb:

    (9) 男の子を撃たれている大砲が灰色です。(Participant 3)

    Otoko-no-ko-o utareteiru taihou-ga haiiro desu.

    Boy-ACC is being shot canyon-NOM grey is.

    (Intended: Boy-ACC shot canyon-NOM grey is. – "The canyon that shot the boy is grey").

2. Use of the nominative instead of the accusative.

    (10) ええと…男の人がとられている…網が茶色です。(Participant 2)

    Eeto... Otoko no hito-ga torareteiru...ami-ga chairo desu.

    Eeh... Man-NOM is catching net-NOM brown is.

    (Intended: Man-ACC is catching net-NOM brown is. – "The net that is catching the man is brown").

(11) ええと…男の子が照らし…照らし…照らしている車です。

(Participant 9)

Eeto... Otoko no hito-ga terashi... terashi... terashiteiru kuruma desu.

Eeh... Man-NOM light...light... is lighting up car is.

(Intended: Man-ACC is lighting up car is. – "(It is) the car that is lighting up the man").

3. Use of the dative instead of the accusative.

(12) ええ…男のほ…人に照ら…照らすのひの人は茶色の制服を着ています。

(Participant 2)

Ee... Otoko no ho... hito-ni tera... terasu-no hi-no hito-wa seifuku-o kiteimasu.

Eeh... Ma... Man-DAT ligh... light up-GEN pers-GEN person-TOP uniform-ACC is wearing.

(Intended: Man-ACC light up person-TOP uniform-ACC is wearing. – "The person who lights up the man is wearing an uniform").

Proportion of the different errors varied with the animacy of the elements, as can be observed in Figures 5.3 and 5.4.

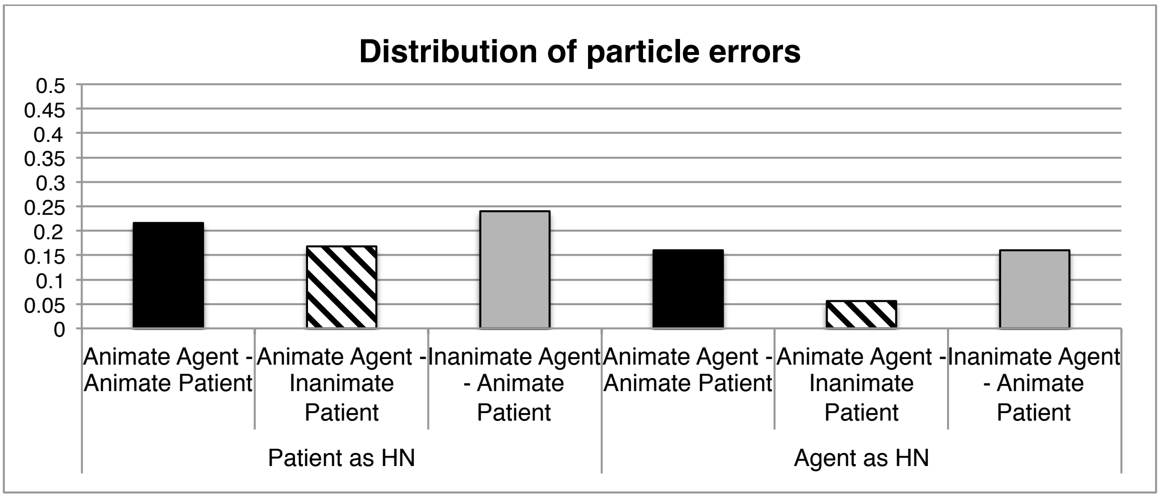


Figure 5.3. Proportion of errors involving changes in case particles by RC type and animacy combination. Proportion was calculated out of the total of 125 errors of changing case particles. 78 of these errors (62.4%) were present in RCs with the patient as HN, while the remaining 47 (37.6%) were commited in the sentences with the agent as HN.

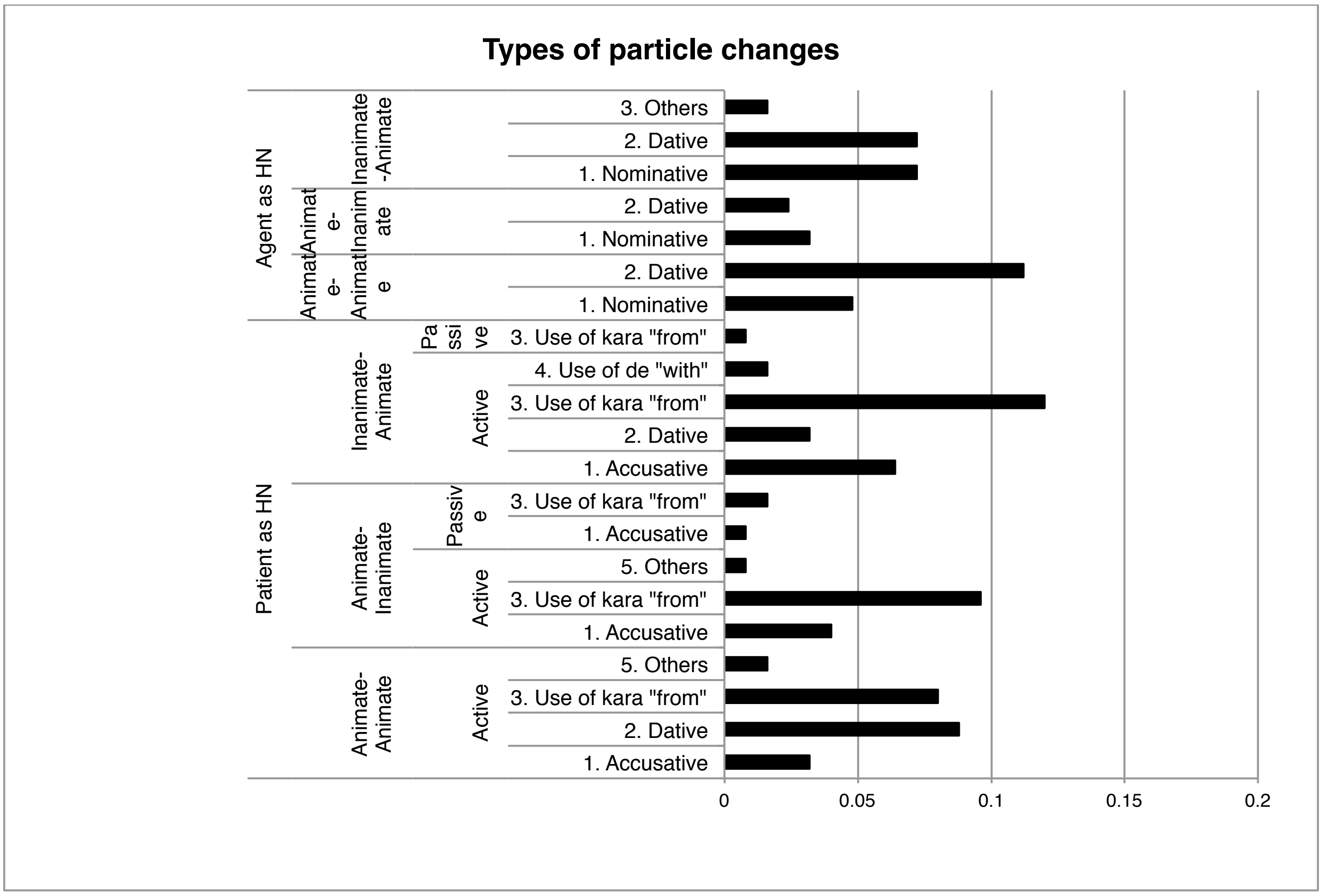


Figure 5.4. Proportion of different types of errors involving change of case particle out of the total number 125. Figures in each category correspond to the list presented above. Note that the scale goes up to 0.2 for better reading.

Errors substituting the dative (which can also be translated as “to”) for the accusative in sentences with the agent as HN are the most frequent in this kind of sentences (20.8% of the total amount of particle change errors; 55.32% of the particle errors found in sentences with the Agent as HN), and are particularly salient when the internal head noun is animate (i.e., in Animate-Animate and Inanimate-Animate sentences): AA sentences: 11.2% of total errors / 29.78% of Agent-HN errors; IA sentences: 7.2% of total errors / 19.15% of Agent-HN errors vs. AI sentences: 2.4% of total errors / 6.4% of Agent-HN errors. This might be interpreted as an influence of the L1. In Spanish, animate direct objects are marked with the word “a”, which translates to “to” in English. This error was present in our 2013 study as well, and shows an overgeneralization of a grammatical rule from the L1.

Another type of error that varies with animacy combination is the change of the accusative to the nominative particle. This error is more prominent in Inanimate-

Animate sentences (7.2% of total errors / 19.15% of Agent-HN errors). This combination is the least prototypical and the one that was more costly for bilingual speakers (see also proportion of valid responses and correct responses in this sentence type). Not only are they semantically atypical, but in this type of sentences, participants must place the animate element at the beginning of the sentence, with the accusative particle, only followed by the verb and the inanimate element, thus reversing semantic prototypicality. To alleviate this burden, participants tended to assign the nominative case to the animate noun, in contrast to the inanimate noun.

In sentences with the patient as HN, most errors involve using the postposition "kara" (from) and using the dative with an active verb (use of kara: 29.6% of total errors / 47.44% of Patient-HN errors; use of dative: 12% of total errors / 19.23% of Patient-HN errors). These errors might come from a failed attempt to construct passive sentences (by using the dative particle without changing the verb) plus an influence of the L1 (by using "kara" –from-, which corresponds to the Spanish preposition "por" (by) used in Spanish to mark passive sentences).

On the other hand, errors consisting of changing the nominative to accusative are again more frequent in the least prototypical Inanimate-Animate events (6.4% of total errors / 10.26% of Patient-HN errors). In this case, the element that has to be placed first is the inanimate, but it must take the role of sentence subject. However, participants seem to make use of semantic cues in this case, by assigning the accusative marker to the inanimate element of the sentence, thus helping them to plan the utterance.

<u>*Pattern of gazes*</u>

Data Analysis. Similarly to previous studies, in order to analyze gaze patterns we followed more stringent inclusion criteria than in the analysis of behavioral data, resulting in a higher number of excluded sentences. In this case, we excluded from the analysis the following responses:

1. Sentences that started with the HN, even if formulated correctly afterwards. Articulation of the HN at the beginning of the sentence will likely confuse the pattern of purely incremental sentence planning. This requirement was also met for monolingual speakers.

2. Sentences with repeated attempts to formulate the correct form (even if this was finally accomplished), or producing the Spanish translation of the sentence before providing the correct response in Japanese.
3. Sentences where the subordinate noun phrase was omitted, failing to overtly express both elements. This requirement was also met for monolingual speakers.
4. Sentences with Speech Onsets with two standard deviations over or below each participant's average.

Both correct and incorrect sentences were included in the analysis, but the analyses with all sentences vs. correct sentences only are presented separately.

Similarly to previous studies, we first grouped gazes to the Area of Interest (AOI) of the agent and to the AOI of the patient in time windows of 50 ms. each. This renders the general pattern of gazes across time. All figures shown in this section indicating the general pattern of gazes are presented with these time windows. Following this, larger time windows were created for statistical purposes. The resulting gaze patterns turned out to be quite complex, making it difficult to spot wide time windows for analysis. Thus, selected time windows were all uniform, with a total of 18 time windows of 400 msec each, so as to ensure measurement after speech onset across all conditions: 0 – 400 ms., 400 – 800 ms., 800 – 1200 ms., 1200 – 1600 ms., 1600 – 2000 ms., 2000 – 2400 ms., 2400 – 2800 ms., 2800 – 3200 ms., 3200 – 3600 ms., 3600 – 4000 ms., 4000 – 4400 ms., 4400 – 4800 ms., 4800 – 5200 ms., 5200 – 5600 ms., 5600 – 6000 ms., 6000 – 6400 ms., 6400 – 6800 ms. and 6800 – 7200 ms.

A separate linear mixed model analysis was then run for each of the time windows, for both gazes to the agent and gazes to the patient as dependent variables separately. The same analysis was repeated twice, first with all the responses, and then only with correct responses. All statistical data are presented in Appendix 4.

The analysis of many smaller time windows provides a more accurate account of the tendencies along the time span. However, it has two major flaws: (1) it complicates the comparison with the results found with monolingual speakers, where wider time windows were used; and, (2) importantly, it might tend to magnify the differences, making small differences statistically significant. In order to undertake fairer comparisons and to clarify whether any observed differences are due to statistical artifacts, we ran another analysis that grouped TWs in a pattern more similar to that used with monolingual speakers. Time windows used for this purpose will be presented below.[45]

For all responses, speech onset was manually calculated using Praat 5.3.71. (Boersma & Weenink, 2014).

Results. Gaze patterns including all responses are shown in Figure 5.5. below. This figure compares gazes for sentences (a) with the Agent as HN, and (b) with the Patient as HN, in all animacy combinations and with responses both in active and passive voice. Speech onset is shown as a vertical line in the graphs. Additionally, speech onset latencies for all conditions are displayed in Table 5.1 below. As can be seen, when the patient was the HN, active forms show longer speech onset latencies than passive forms or than responses with the agent as HN. Concerning animacy, the AI condition was the fastest, with AA and IA conditions showing similar latencies in both sentences with agent as the HN and sentences with patient as the HN.

	Agent as HN	Patient as HN			General
		Active	Passive	General	
Animate agent - Animate patient	5011.03	7921.40	4240.6	4766.43	4923.97
Animate agent - Inanimate patient	4632.03	4733.57	4280.45	4507.01	4555.44
Inanimate agent - Animate patient	4940.96	8703.26	3928.17	4610.32	4769.52
General	4814.42	5712.12	4130.09	4602.88	4713.28

Table 5.1. Speech onset latencies in milliseconds for all conditions (correct responses only).

The general analysis revealed that there was a main effect of RC from TW1 (0-400 ms.), for both agent and patient. From 0 to 800 ms. (TW1 and TW2), gazes were directed to a larger extent to the agent when the sentence had the agent as HN, and to the patient when this was the HN. This pattern was especially acute for the 400

[45] As it will be seen below, the first time window selected for comparison is shorter than 400 ms. and it is intended for comparison between studies. The other three windows are of varying lengths and their main purpose is to corroborate found trends from a more statistically reliable period of time. They group the period of times in which a determined trend is visible.

to 800 ms. TW (TW1: gazes to agent: Agent-HN = 0.20, Patient-HN = 0.12, t = 2.88, $p = 0.005$; gazes to patient: Agent-HN = 0.08, Patient-HN = 0.17, t = -3.547, $p < 0.001$; TW2: gazes to agent: Agent-HN = 0.38, Patient-HN = 0.16, t =3.834, $p = 0.002$; gazes to patient: Agent-HN = 0.18, Patient-HN = 0.38, t = -4.035, $p < 0.001$). From 850 ms. to 1600 ms. a crossing point is apparent. Statistically, there are no differences in the proportion of gazes to agent and patient in either sentence with the agent as HN and with the patient as HN (all ts < 2, although main effect of RC in gazes to agent is marginally significant in TW3: t = 1.89; p = 0.09). After this transition period, from 1600 to 3200 ms. (TWs 5 to 8) participants started looking more extensively at the element they were going to utter first, which was not the head noun: proportion of gazes to the agent was greater in sentences with the patient as HN (gazes to agent: TW5: agent-HN = 0.22, patient-HN = 0.31; TW6: agent-HN = 0.25, patient-HN = 0.39; TW7: agent-HN = 0.29, patient-HN = 0.36; TW8: agent-HN = 0.25, patient-HN = 0.40) and, correspondingly, gazes to the patient were greater in sentences with the agent as HN (gazes to patient: TW5: agent-HN = 0.34, patient-HN = 0.22; TW6: agent-HN = 0.34, patient-HN = 0.19; TW7: agent-HN = 0.33, patient-HN = 0.21; TW8: agent-HN = 0.38, patient-HN = 0.19) (all differences $p < 0.05$, except for Patient in the 1600-2000 TW ($p = 0.057$), and Agent in the 2400-2800 TW ($p = 0.056$), which were only marginally significant).

From 3200 to 4800 ms. there was an extended period in which there were no differences in proportion of gazes due to type of RC. As can be seen in Figure 5.5, gazes to both agent and patient are very close to each other until 4800 ms. From this moment onwards, beyond speech onset, participants are looking at the element they are going to utter in the second place to a larger extent, namely the agent, in sentences with the agent as HN, and the patient, in sentences with the patient as HN (TW4800-5200: gazes to agent: t = 2.769, $p = 0.012$, gazes to patient: t = -2.66, $p = 0.017$; TW5200-5600: gazes to agent: t = 2.695, $p = 0.0235$; TW6000-6400: gazes to patient: t = -4.815, $p < 0.0001$; TW6400-6800: gazes to agent: t = 2.93, $p = 0.012$, gazes to patient: t = -5.448, $p < 0.0001$; TW6800-7200: gazes to agent: t = 5.273, $p < 0.0001$, gazes to patient: t = -4.335, $p = 0.0015$; however from 5200 to 6000, along with gazes to patient from 5200 to 5600, and to agent from 6000 to 6400 ms., differences are only marginally significant). Interestingly, in RCs with patient as the HN, speech onset is delayed until 5810 ms., in contrast to the 4996 ms. onset for sentences with the agent as HN. That is, especially in sentences with patient as the HN, it seems that gazes are fixated significantly to a greater extent on the second element before articulation begins.

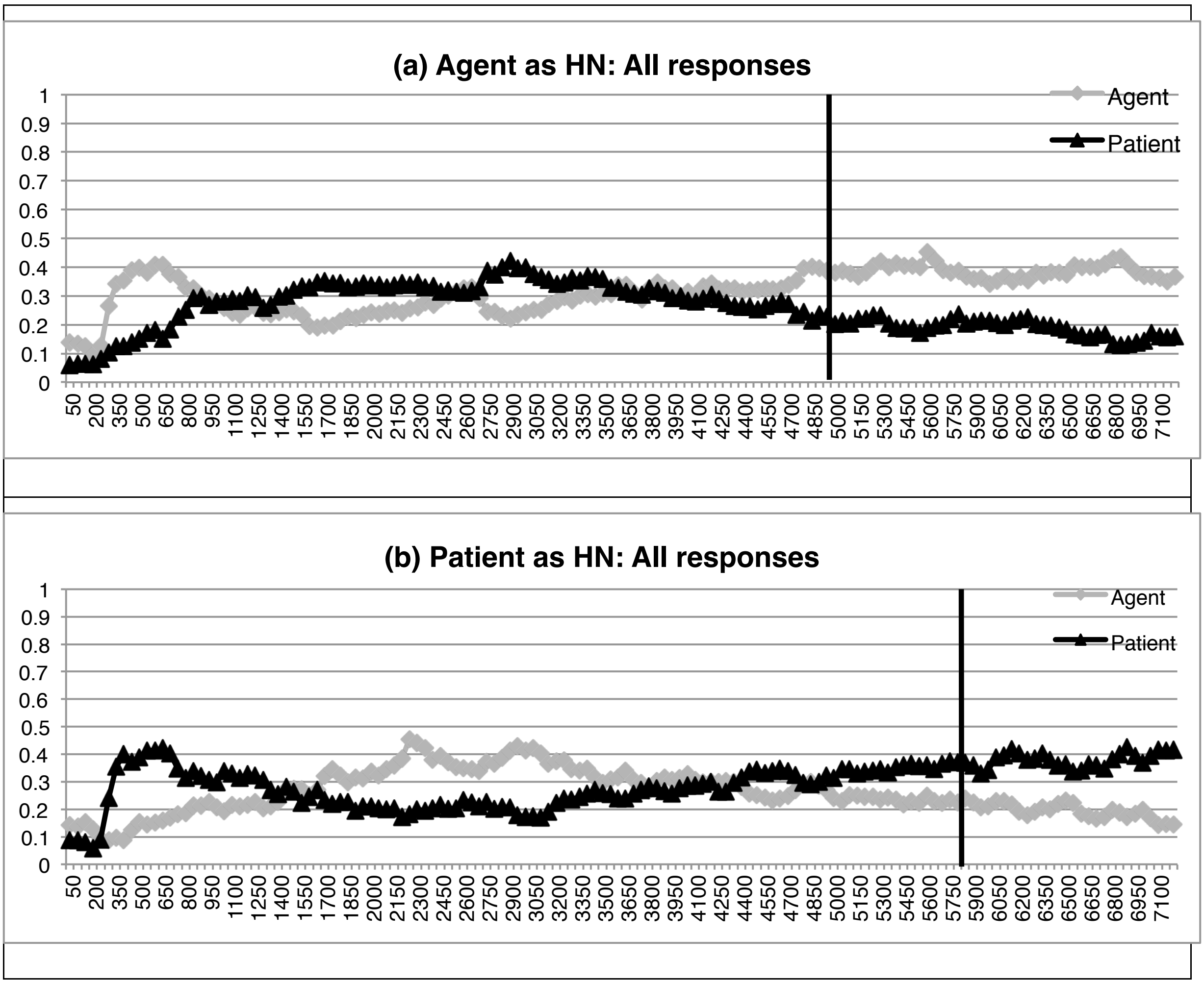


Figure 5.5. Gaze patterns to Agent and Patient in sentences with the Agent as HN (a), and sentences with the Patient as HN (b). All animacy combinations have been collapsed and both correct and incorrect sentences are included. Vertical lines indicate speech onset.

On the other hand, animacy did not exert any significant effect along the time course, that is, in any of the 18 analyzed TWs.

When only correct responses were taken into account (Figure 5.6), the landscape was very similar. A main effect of RC type was found for gazes to the agent and to the patient from 0 to 800 ms. (TW1: gazes to agent: Agent-HN = 0.19, Patient-HN = 0.11, t = 1.98, $p = 0.053$, gazes to patient: Agent-HN = 0.10, Patient-HN = 0.17, t = -2.273, $p = 0.026$; TW2: gazes to agent: Agent-HN = 0.39, Patient-HN = 0.20, t = 4.407, $p < 0.0001$, gazes to patient: Agent-HN = 0.20, Patient-HN = 0.37, t = -2.341, $p = 0.035$): gazes were directed mainly to the HN of the sentence, which is not the first mentioned element in Japanese. After a period in which gazes to agent and patient cross and the pattern reverses, we found again a main effect of RC type from 1600 to 3200 ms. (TWs 5 to 8) (TW5: gazes to agent: t = -4.075, $p < 0.0001$, gazes to patient: t = 4.647, $p < 0.0001$; TW6: gazes to agent: t = -4.656, $p < 0.0001$, gazes to patient: t = 2.604, $p = 0.035$; TW7: gazes to agent: t = -2.069, $p = 0.04$, gazes to patient: t = 2.643, $p = 0.009$; TW8: gazes to agent: t = -3.113, $p = 0.002$, gazes to patient: t = 3.536, $p = 0.0006$). In this case, participants fixated extensively on the element they were going to mention first, that is, the subordinate element of the relative clause. Following a period in which there are no differences due to RC type (both gazes to agent and patient are close to each other), from 4800 to 7200 ms. participants look to a larger extent at the HN[46], which is the element they are going to produce in the second place. This effect is visible when analyzing only correct responses (TW4800-5200: gazes to agent: t = 3.513, $p = 0.0009$, gazes to patient: t = -2.929, $p = 0.0048$; TW5200-5600: gazes to agent: t = 5.261, $p < 0.0001$, but gazes to patient: t = -1.802, $p = 0.12$ (n.s.); TW5600-6000: gazes to agent: t = 5.013, $p < 0.0001$, gazes to patient: t = -5.454, $p < 0.0001$; TW6000-6400: gazes to agent: t = 5.393, $p < 0.0001$, gazes to patient: t = -6.125, $p < 0.0001$; TW6400-6800: gazes to agent: t = 3.289, $p = 0.013$, gazes to patient: t = -5.5, $p < 0.0001$; TW6800-7200: gazes to agent: t = 4.096, $p < 0.0001$, gazes to patient: t = -5.989, $p < 0.0001$) and also when taking into account all responses, although in that case differences from 5200 to 6000 are only marginally significant, as well as differences in gazes to patient from 5200 to 5600, and to agent from 6000 to 6400 ms. (as for the rest of differences: TW4800-5200: gazes to agent: t = 2.769, $p = 0.012$, gazes to patient: t = -2.66, $p = 0.017$; TW5200-5600: gazes to agent: t = 2.695, $p = 0.0235$; TW6000-6400: gazes to patient: t = -4.815, $p < 0.0001$; TW6400-6800: gazes to agent: t = 2.93, $p = 0.012$, gazes to patient: t = -5.448, $p < 0.0001$; TW6800-7200: gazes to agent: t = 5.273, $p < 0.0001$, gazes to patient: t = -4.335, $p = 0.0015$).

[46] As can be seen in Appendix 4, the main effect of RC type was marginally significant for gazes to the patient from 4400 to 4800 ms. ($p = 0.068$). Differences in gazes to the agent in this TW are not significant ($p = 0.14$), but they are numerically different (Patient-HN: 0.29 vs. Agent-HN: 0.38). This might suggest an earlier start of the planning of the second element in correct responses than when taking all sentences into account.

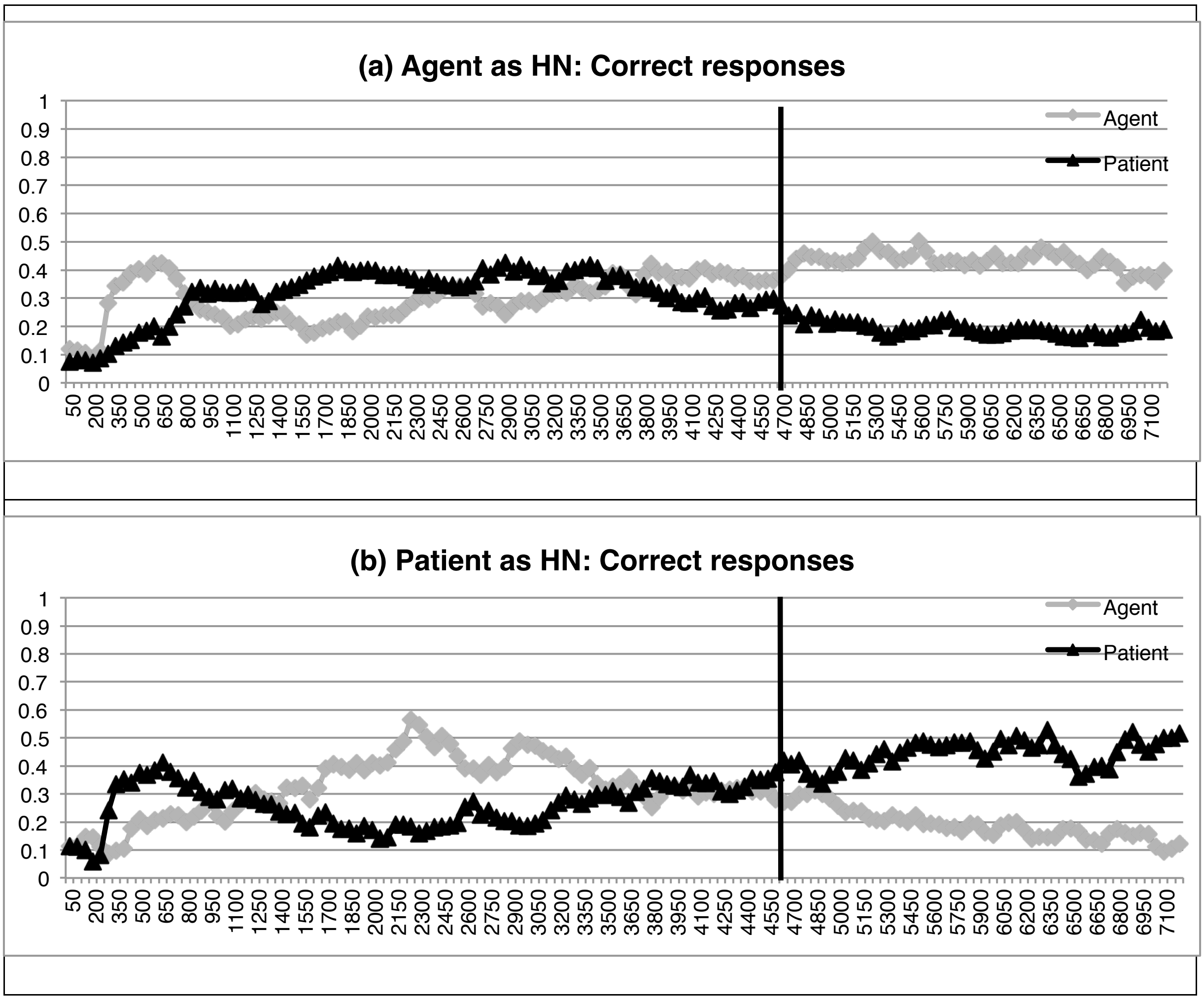


Figure 5.6. Gaze pattern to Agent and Patient in sentences with Agent as the HN (a) and sentences with Patient as the HN (b). With all animacy combinations collapsed and including only correct sentences. Vertical lines indicate speech onset.

In contrast to the global analysis, when analyzing only the correct responses we found some effects of animacy in gazes to agent and patient from 1600 to 2000 ms. (gazes to agent: t = 2.712, p = 0.009; gazes to patient: t = -2.323, p = 0.03). In this TW, gazes to the agent were higher in the AI and the IA conditions (0.34 and 0.32, respectively) than in the AA condition (0.16), while gazes to the patient were higher in the AA condition (0.47) than in AI and IA conditions (0.26 and 0.27, respectively). Moreover, there was a marginal effect of animacy in gazes to the patient from 0 to 400 ms. (t = 1.872, p = 0.07), with participants focusing more extensively on the patient in the IA condition (0.22) than in the AA (0.11) and AI conditions (0.08). In Figures 5.7 and 5.8, patterns of gazes to agent and to patient along the time span across the different animacy combinations are shown.

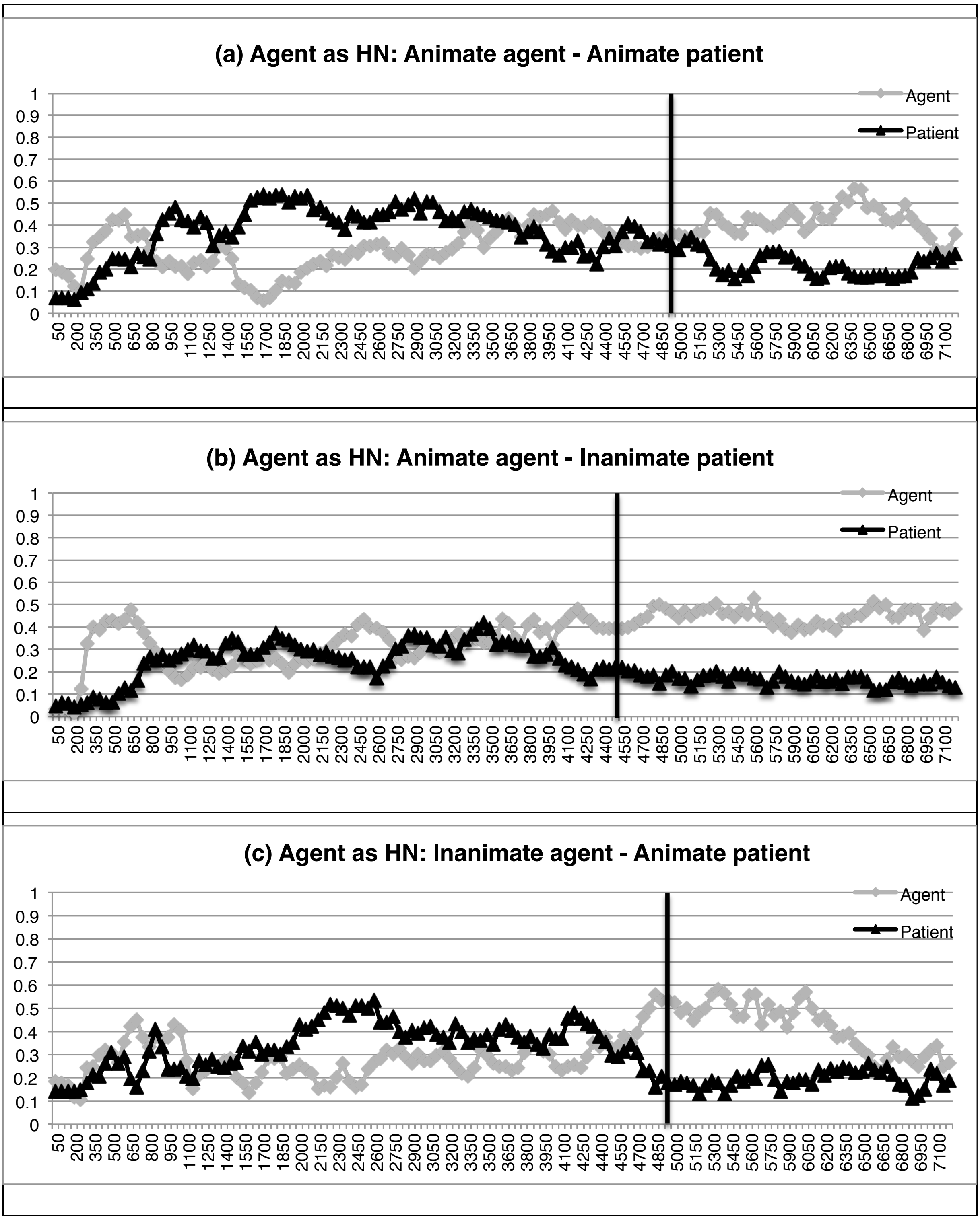


Figure 5.7. Gazes to agent and patient across different animacy combinations in sentences with the agent as HN. Only correct sentences are included. Vertical line indicates speech onset.

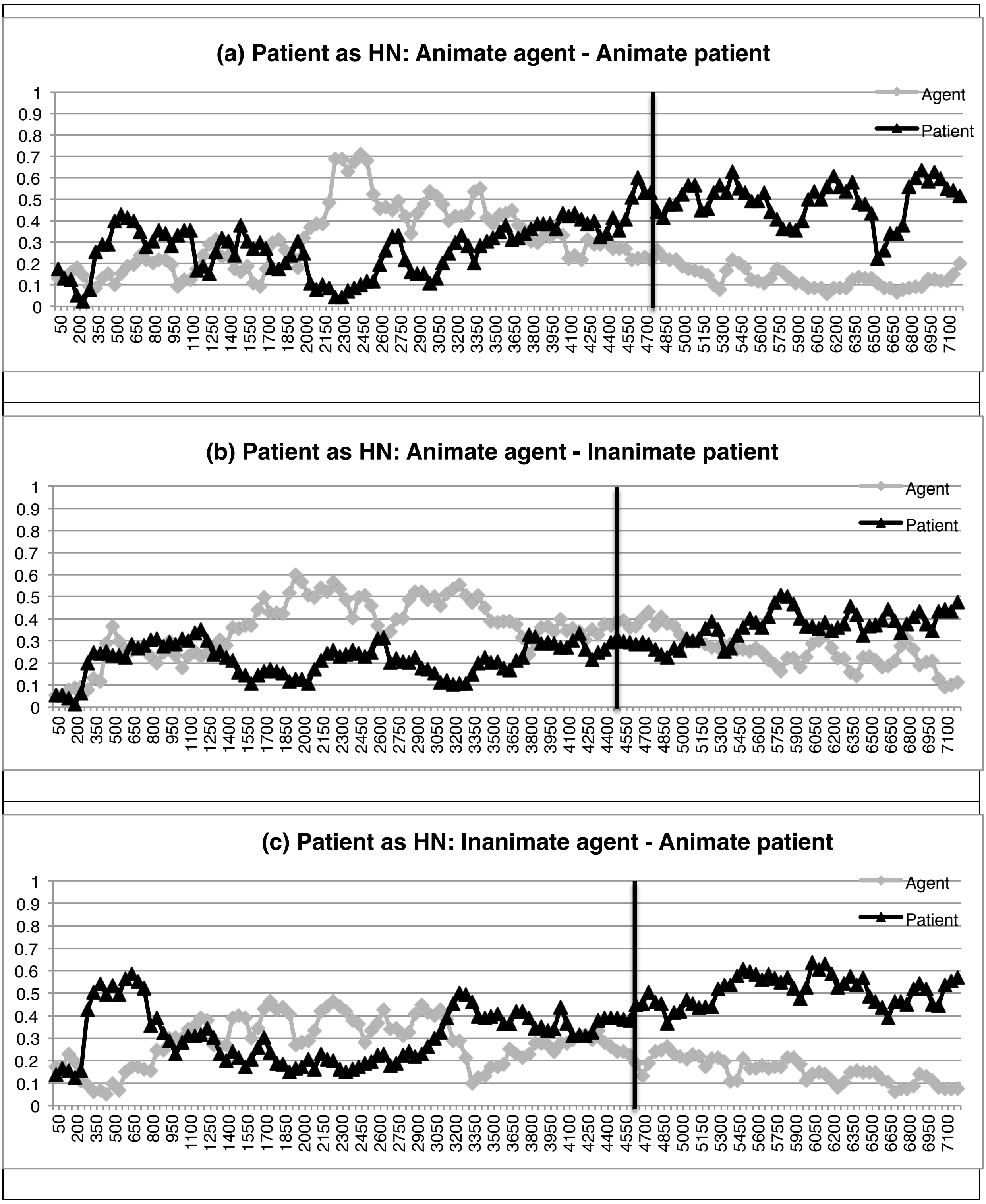


Figure 5.8. Gazes to agent and patient across different animacy combinations in sentences with the patient as HN. Only correct sentences are included. Vertical line indicates speech onset.

Subsequently, we carried out an analysis with wider time windows and/or time windows comparable to those of native speakers, in order to allow for more direct comparison between experiments in the three main issues under analysis. First, similarly to the case of monolinguals, there was an effect of RC-type from TW1 onwards in the case of bilinguals. However, the TW used here was wider (for the sake of keeping uniformity along the analysis). In our studies with monolinguals, we used a TW from 0 to 350 ms.. In this case, fixations on the HN started from 300 ms. onwards., although we found differences between Spanish and Japanese speakers (with Spanish speakers focusing slightly later on the HN). In order to find out whether exactly the same timing is present in bilinguals, we ran another analysis, this time with a time window spanning 0 to 350 ms. only (issue 1 above). Despite the reduction of the TW, when all responses were taken into account there was an effect of RC for gazes to the patient (t = -2.558, $p = 0.012$) and marginally to the agent (t = 1.953, $p = 0.053$). Considering only correct responses, however, there was no effect in gazes to agent (t = 1.225, $p = 0.22$) and marginal effects in gazes to the patient (t = -1.978, $p = 0.076$). In addition, there was no effect of animacy in this TW when taking into consideration all items, and only a marginal effect of gazes to the patient when considering only correct responses (t = 1.886, $p = 0.07$). Similarly as before, there were more gazes directed to the patient in IA sentences (0.20) than in AA (0.17) and AI (0.12) sentences. When analyzing only RCs with the patient as HN, we found an interaction between animacy and voice, which was significant when including all responses in the analysis (t = 2.125, $p = 0.035$) and only marginally significant with correct responses only (t = 1.855, $p = 0.067$). Numerically, passive sentences, as well as IA sentences, tend to gather more gazes in general. However, the pattern is not the same for all animacy combinations: AA sentences present the opposite pattern with more gazes directed to the patient when the sentence is going to be active (that is, when the patient is *not* going to be the subject). Moreover, the larger proportion of gazes to the patient in passive sentences in comparison with active sentences is enhanced in the IA condition in comparison with the AI condition (a greater tendency to focus on the patient in passives –i.e. the element that is going to be the subject– when it is animate than when it is inanimate).

Secondly, we analyzed the TW from 400 ms to 1000 ms., where we had found Japanese monolinguals to fix their initial gazes on the HN (issue 2). We wanted to analyze whether the same effect within the same time window would show up in bilingual speakers as well. In this TW, there was a main effect of RC type for gazes to agent and patient in bilinguals, both when taking into account all responses (gazes to agent: t = 3.739, $p = 0.003$; gazes to patient: t = -4.932, $p < 0.0001$) and with correct responses only (gazes to agent: t = 4.124, $p < 0.0001$; gazes to patient: t = -2.58, $p = 0.012$).

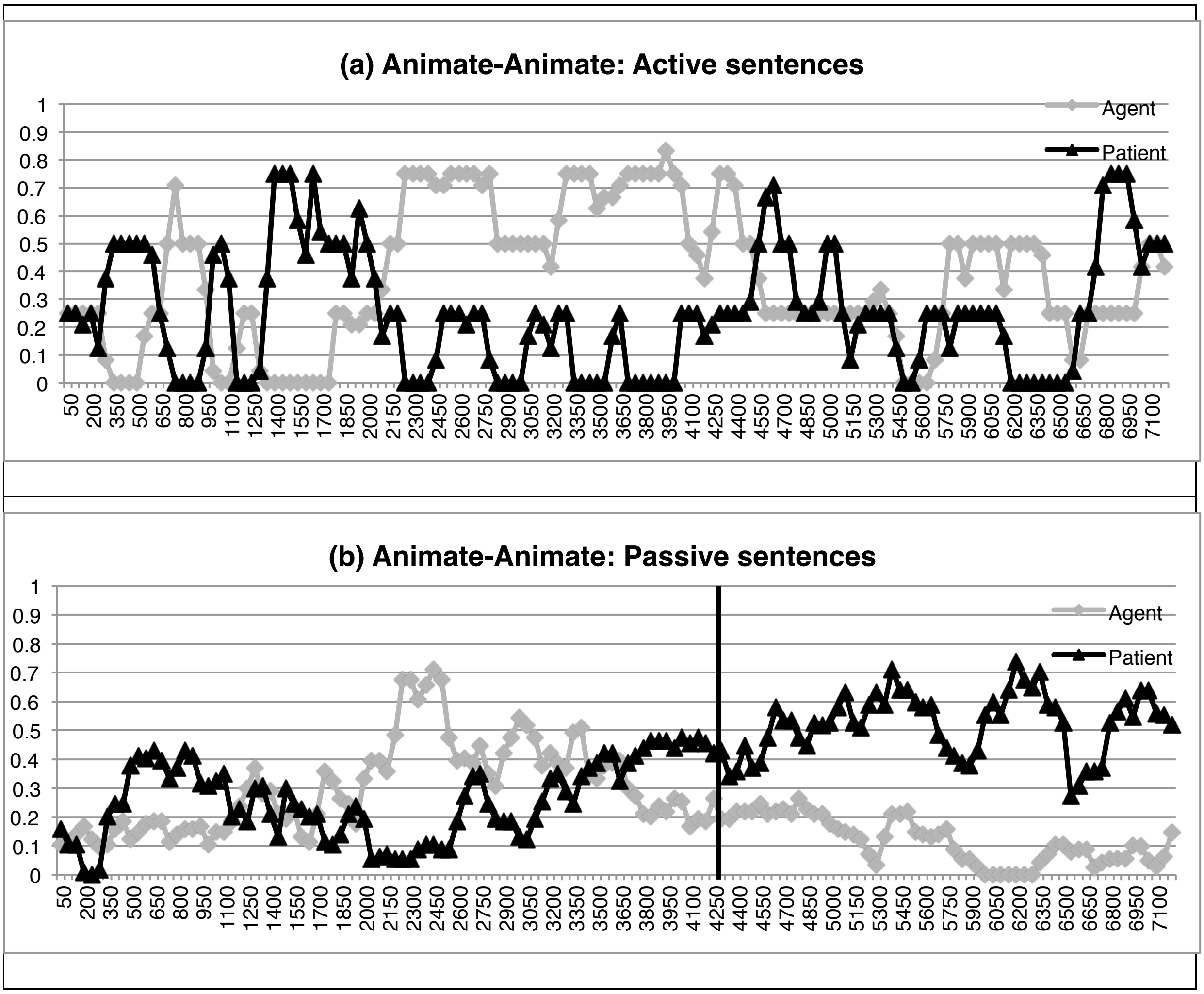
Subsequently, we analyzed larger TWs encompassing the smaller TWs in which significant trends were found, so as to test whether the same trends appear with longer time spans. To this end, we chose three time windows: from 1500 to 3200 ms., from 3200 to 4800 ms. and from 4800 to 7200 ms.

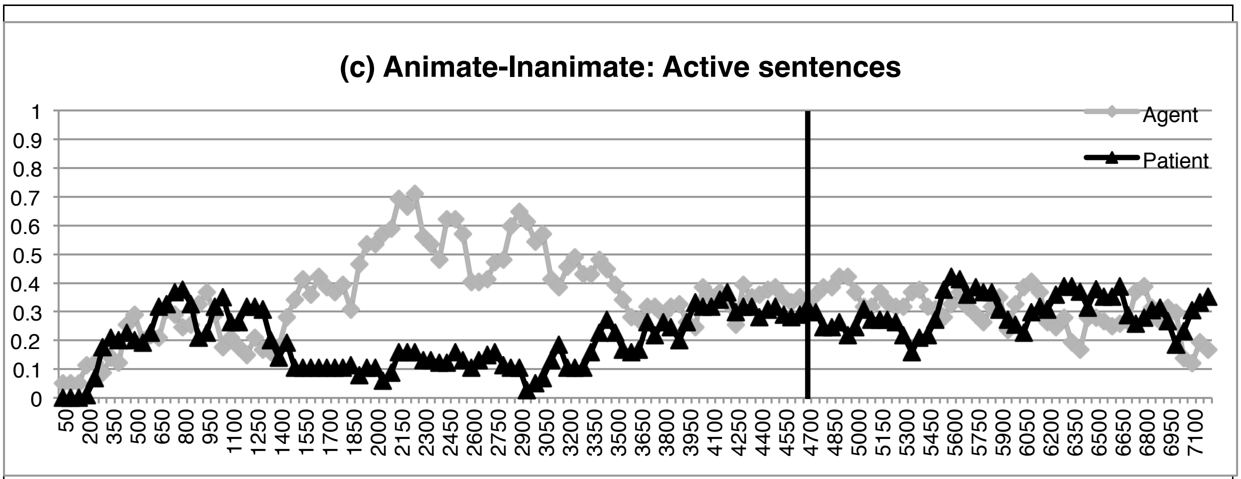
The trends found with smaller time windows were replicated: from 1500 to 3200 ms.: there is a main effect of RC type for both gazes to agent and patient, and both in all sentences (gazes to agent: t = -2.772, $p = 0.015$; gazes to patient: t = 2.671, $p = 0.022$) and in correct sentences only (gazes to agent: t = -5.338, $p < 0.0001$; gazes to patient: t = 5.431, $p < 0.0001$), showing a trend to direct gazes to the element that is going to be uttered first. In the next TW, right before speech onset (from 3200 to 4800 ms.), there were no differences due to RC type, with participants focusing equally on agent and patient during this time interval (all ts < 2). Finally, after speech onset (from 4800 to 7200 ms.) we found main effects of RC type again. In this case, participants focused on the element they were going to articulate in the second place (the HN of the relative clause): all responses - gazes to agent: t = 2.565, $p = 0.023$; gazes to patient: t = -2.565, $p = 0.011$; correct responses - gazes to agent: t = -2.594, $p = 0.048$; gazes to patient: t = -2.862, $p = 0.06$.

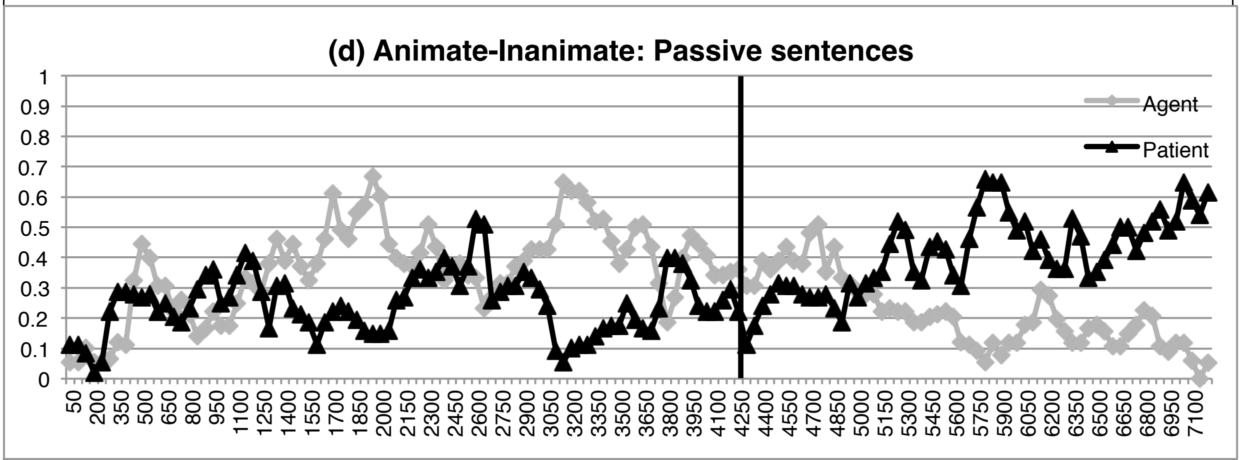
After the first analysis performed on all RC types, we ran another one only with sentences with the Patient as HN, so as to address the effects of the choice of voice on gaze patterns (i.e. gaze patterns in sentences in active vs passive voice) (issue 3). In this comparison we analyzed once again data from all responses and only from correct responses. It is important to note that the amount of data included in this analysis is considerably scarce, with 169 utterances in total, of which only 89 utterances are correct: 23 animate-animate sentences (4 actives and 19 passives), 37 animate-inanimate sentences (19 actives and 18 passives) and 29 inanimate-animate sentences (4 actives and 25 passives). Note that the number of active sentences is especially low in the AA and IA conditions, given the tendency to produce passive sentences under these two animacy conditions (see behavioral data results above, for more details). In general, the avoidance of active forms under these two animacy conditions led to longer speech onsets, almost doubling the time of the corresponding passive sentences in the same animacy combinations: SOs for the AA condition for actives – 7921 ms., for passives – 4241 ms.; SOs for the IA condition: for actives – 8703 ms., for passives – 3928 ms. This will be reflected in the graphs we will present below (see Figure 5.9), in which the data corresponding to these categories are not clear. Similarly, statistical analyses are tentative, and suggestive of possible tendencies to be further explored.

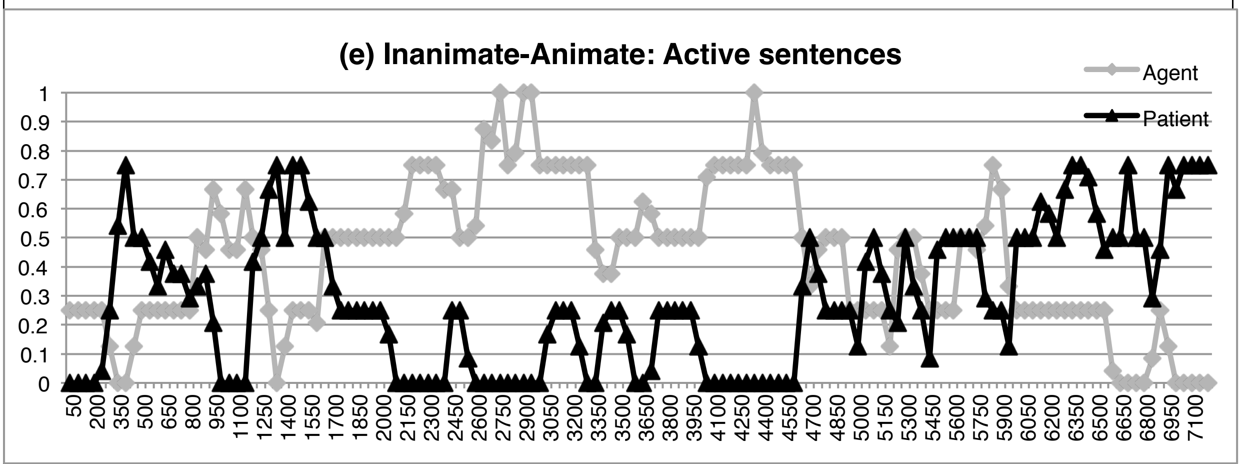
Similarly to previous analyses, we ran different LME analyses for both gazes to the agent and to the patient in 400 ms. time windows, from 0 ms. (picture onset) to 7200 ms., the latter beyond the speech onset. The results of the analyses showed no effects of animacy or voice from 0 to 2000 ms. After this period, from 2000 to 2800 ms. (TWs 6 and 7) there is a main effect of voice on gazes to the agent when taking into account only correct responses (TW6: $t = -2.134$, $p = 0.036$; TW7: $t = -2.422$, $p = 0.023$) and a marginal effect on gazes to the agent from 2800-3200 ($t = -1.911$, $p = 0.06$), a period in which the effect on gazes to the patient was significant ($t = 2.281$, $p = 0.02$). Thus, during this period participants looked more extensively to the agent when the response was going to be in the active form, and to the patient in passive form, though to a lesser extent**.** When taking into account all responses, the pattern differs, showing an effect of voice in gazes to the agent from 1200 to 1600 ms. ($t = 2.885$, $p = 0.006$), and from 2000 to 2400 ms. ($t = -2.463$, $p = 0.025$). However, in this case, there is a tendency to focus more extensively to both agent and patient when the sentence is going to be in passive form, showing greater planning costs in general for these structures for bilinguals with greater amount of incorrect responses. After this period, all differences fade away until 5200 ms., except for gazes to patient in the 3600-4000 TW with all responses included ($t = 2.575$, $p = 0.022$), where the same pattern that was earlier reported with correct responses appears, namely, after speech onset for passive sentences (and active ones in the AI condition), but before speech onset for active sentences in the AA and IA conditions. From this moment onwards (until the end of the timeline at 7200 ms.), there is again a main effect of voice on responses when analyzing all responses and when analyzing only correct responses, with participants directing more gazes to the agent when the sentence was in the active form and more gazes to the patient when it was in the passive form (all responses: patient 4800-5200: $t = 2.431$, $p = 0.044$, patient 5200-5600: $t = 3.028$, $p = 0.006$; agent 5600-6000: $t = -2.784$, $p = 0.007$; patient 5600-6000: $t = 3.501$, $p =$

0.0015; agent 6000-6400: t = -2.563, p = 0.013; patient 6000-6400: t = 3.867, $p <$ 0.0001; correct responses: patient 5200-5600: t = 3.124, p = 0.0026; agent 5600-6000: t = -3.59, $p <$ 0.0001; patient 5600-6000: t = 2.209, p = 0.029; agent 6000-6400: t = -3.217, p = 0.0029; patient 6000-6400: t = -2.147, p = 0.034; agent 6400-6800: t = -2.147, p = 0.034; agent 6800-7200: t = -2.18, p = 0.032). As shown in Figure 5.9, this pattern somewhat varied with animacy, not only from 0 to 350 ms., as presented before, but also from 6000 to 6400, where there was an interaction between animacy and voice for gazes to the patient when taking into account only correct responses (t = -2.169, p = 0.032), a difference that was marginal for the agent when taking into account all responses (t = 1.799, p = 0.073). In contrast to what we found in the first TW, here we found that the general gaze pattern shows up in the AA and AI conditions (i.e. more gazes to the patient when the sentence was in passive form), whereas the opposite pattern appears in the more atypical IA sentences. General gaze patterns found in active and passive forms for all animacy combinations can be seen in Figure 5.9.




(c) Animate-Inanimate: Active sentences
Agent
Patient


(d) Animate-Inanimate: Passive sentences
Agent
Patient


(e) Inanimate-Animate: Active sentences
Agent
Patient

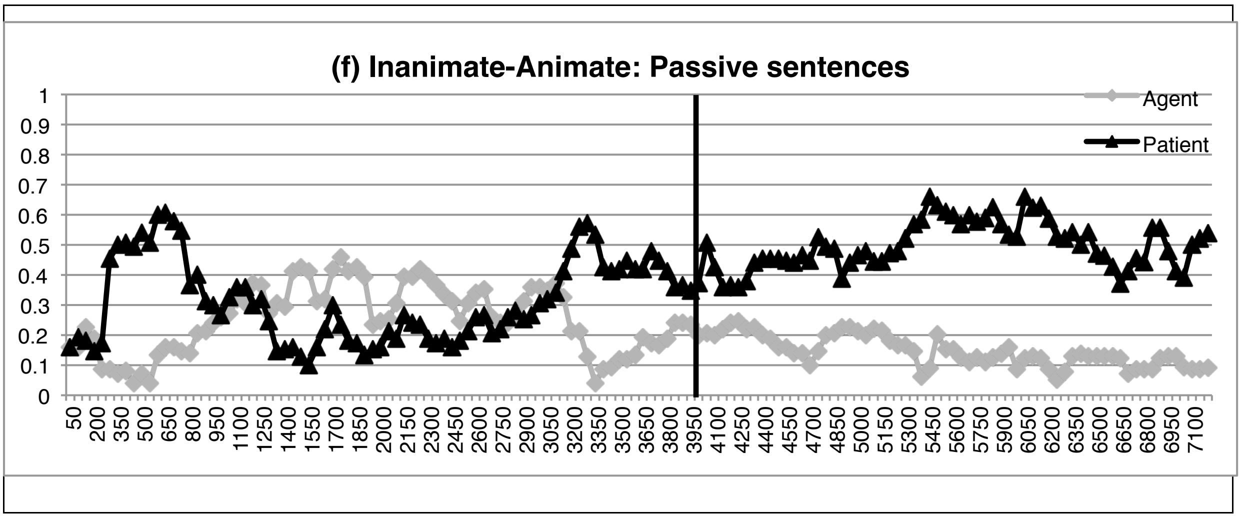


Figure 5.9. Gaze patterns to Agent and Patient in sentences with the Patient as HN across the three animacy conditions (Animate-Animate, Animate-Inanimate and Inanimate-Animate) in active and passive forms. Data include only correct sentences. Vertical lines indicate speech onset, except for the Active forms of AA and IA conditions, where speech onset was beyond the measured 7200 ms.

Discussion

In this study, we explored the production planning of RCs by Spanish-Japanese bilinguals. These speakers are faced with the task of producing a structure with the opposite word order to that of their native language, placing the HN at the end of the clause. Word order is usually not a problem for bilingual speakers, who learn it fairly quickly in their acquisition process. However, data concerning online production of RCs were still lacking. In this study we aimed to analyze the effects of animacy and word order in planning, focusing on three temporal stages that were found to be relevant for monolingual speakers: apprehension, grammatical encoding and lexical access. We will first discuss the overall results (both behavioral and eye-tracking data) for this group of participants, relating it subsequently to the results obtained for Spanish and Japanese monolingual speakers. We will conclude with the implications of our results for incremental production models.

An analysis of the spoken utterances by bilingual speakers reflected they were affected by animacy in their responses: the only unambiguous and most prototypical category (animate agent – inanimate patient) was the one that yielded fewer errors. Similarly, particle errors that were not based on the L1 were mainly related to the animacy of the elements involved: a tendency to assign the nominative particle to the animate element and the accusative particle to the inanimate one. Not only errors had an influence of animacy, but also the preferred voice of responses. When the patient was the HN of the RC, bilingual speakers preferred to produce AA and IA sentences as passives, thus making the animate element and HN the subject of the sentence. The proportion was more similar to Japanese speakers than to Spanish native speakers, despite the higher grammatical difficulty of this structure. This might reflect the high visibility of the tendency Japanese speakers have to use passive sentences, which could be used strategically when learning Japanese, even at the risk of overuse. However, it could also reflect a tendency of bilingual speakers to overly rely on animacy patterns to minimize planning difficulties. In other words, in the case of AA sentences, assigning the animate HN the subject role would help them reduce any extra cognitive effort coming from the competition of the two nouns (Gennari *et al.,* 2012). Similarly, in the IA condition, assigning the subject function to the animate element and the by-phrase function to the inanimate element helps to ease planning of the different lexical items and their relations. Whether this is a conscious learning strategy or just part of the incremental planning process in bilingual speakers might become clearer by exploring gaze patterns.

Perhaps the most remarkable fact of the performance of our bilingual group is that the utterances they produced and the underlying planning processes have turned out to be more similar to what we found in native speakers of Japanese than to the pattern of Spanish monolingual speakers. In both bilingual and Japanese monolingual speakers we found a pattern in which speakers fixate the HN (from 300-350 to 1000 ms., with the same length in both groups, the only difference being an earlier fixation on the HN in the case of monolinguals) before shifting to the element they were going to articulate in the first place (the internal NP). It is also noteworthy that speech onset latencies did not differ between Japanese monolinguals and bilingual speakers, both being longer than Spanish monolingual speakers. However, there are also some differences along the process between native Japanese and bilingual speakers that we would like to draw our attention to in each of the three planning stages we have emphasized in the current study:

1. *Apprehension stage:* in this study, we analyzed the pattern of fixations to agent and patient during the first milliseconds of exposure to the picture. In the case of bilingual speakers, as it was the case in monolinguals, speakers locate and start fixating the HN within the first 350 ms. (300 ms. for Japanese

native speakers, 350 ms. for bilingual speakers), with a sharp increase of gazes from 350 ms. to 400 ms. Gazes are equally distributed between both participants initially, which purports to show that speakers are trying to secure a fast and overall understanding of the scene in front of them before focusing on the element they have to say something about (the HN of the sentence, of which something is being asked: e.g. the person who is wearing a pink blouse). However, in the case of RCs with the patient as the HN, there was an interaction between animacy and voice during the first 350 ms. in gazes to the patient. This interaction is difficult to explain: although in general participants focused more extensively on the patient when they were going to utter a passive sentence, this pattern was reversed in AA sentences. On the other hand, bilingual speakers tend to fixate the patient more in active sentences, that is, where the patient is not the subject but remains as the object, despite being the HN. This trend makes it difficult to ascertain whether there is an influence of animacy on early TWs, as predicted on the assumption that bilingual speakers in general rely more strongly on animacy cues to help them process and plan speech in their non-native language (MacWhinney, 1997; Sasaki, 1994), or whether this difference is due to greater difficulty in planning the speech (i.e. learners face the task of creating a sentence in which the most dominant element will not be the subject). Obviously, we need more data to clarify this interaction, since many different factors may influence the early stages of bilingual sentence planning.

2. *Grammatical encoding:* as was mentioned at the beginning of this discussion, a major result from this study was the finding that bilingual speakers planned their utterance in a very similar fashion to Japanese monolingual speakers, which happens to be far removed from the way speakers of their own native language do. These speakers start focusing on the HN from 400 to 1000 ms. before shifting to the element they are going to utter first. Despite their lower proficiency in Japanese, they do not take longer than native speakers, which suggests that this process did not demand an extra cognitive effort to our group of bilingual speakers. Remember that in the case of Spanish, monolingual speakers started looking at the HN –the first mentioned NP– within the first 400 ms. and kept looking at this element until speech onset.

This period was thought to include both structural and lexical planning in a presumably interwoven and closely related way. Thus, Spanish speakers made use of a more lexical, incremental planning than their Japanese counterparts did, while bilingual speakers apparently rely more heavily on structural planning before attempting lexical retrieval. Two possible conclusions may be drawn from these results: first, it might be the case that bilingual speakers are able to change their planning strategy to accommodate the requirements of their L2. Speech planning is characterized by a largely flexible system, which could be adapted to the structure on hand (Norcliffe & Konopka, 2105). Thus, it might not be surprising to assume that speakers are able as well not only to adjust their strategies *within* a language, but also *between* two or more languages, provided that they are sufficiently proficient in their L2. However, we cannot entirely discard the notion that monolingual Spanish speakers are also reliant on structural planning (as the hierarchical incrementality model would claim) in their native language. As mentioned in the previous chapter, the SVO canonical structure dominant in Spanish makes it difficult to tell apart linear and hierarchical incrementality accounts of sentence planning.

3. *Planning during lexical access:* Bilingual speakers showed a production planning that appears to be highly incremental from the beginning of their processing. Complex sentences like relative clauses did not stem incremental planning, encouraging a conscious strategy of planning all the elements in the sentence beforehand. Rather, processing was remarkably similar to that of native speakers from the very begininng. Similarly, animacy did not have a strong influence, but only confined to a brief temporal period. This period corresponds to the moment at which speakers are fixating the first element to be uttered. However, even though animacy itself was not a guiding factor in planning, gaze patterns show at least a tendency for animate elements to be more fixated than inanimate ones. This results in increased or decreased differences between agent and patient, depending on the animacy combinations and the structure that is being planned, as shown in Figures 5.7 and 5.8 above. This pattern is, once again, very similar to the one found in monolingual Japanese speakers.

However, there was a major difference between monolingual and bilingual speakers as regards the ease of lexical access. For bilingual speakers, there was a period of approximately 1300 ms. from lexical access to speech onset. During this period, participants were looking equally at both agent and patient. Once again, this could be due to two possible scenarios. One possibility is that there is an overlapping activation of the lemmas of the agent and the patient: after the retrieval of the first element, and due to a certain amount of activation of the HN (second element), some degree of lexical competition may ensue (Gennari, *et al.,* 2012). As Gennari, *et al.* suggested, this competition can be the main factor that leads to a higher number of passive sentences in the AA condition than in the AI condition[47]. This would mean that overlapping activation occurs for both monolingual and bilingual speakers. However, monolinguals are more skillful in resolving this competition and starting speech right over. Bilinguals, in turn, would face greater difficulty in resolving this conflict. On the other hand, a second possibility is that speakers are actually planning the verb during this time window. Previous work has suggested that the periods during which agent and patient are equally fixated reflect planning of the action (see , Norcliffe *et al.*, 2015; Norcliffe & Konopka, 2015; Griffin & Bock, 2000), and this could be the underlying reason for lexical competition (remember that the verb is the second uttered element in Japanese RCs). For Japanese monolinguals, verb planning takes a very short time (especially in the current task, since the verb was provided in advance). In contrast, for bilingual speakers planning of the verbal form and its morphology could have been harder, thus making them linger in that phase for a longer time. With our present data it is difficult to determine which interpretation is more adequate. Moreover, competition between lexical items would result in a longer delay for AA sentences than AI or even IA sentences. Although statistically this seems not be the case, it cannot be entirely discarded due to the complex relations between lexical

[47] The fact that IA sentences also showed a high proportion of passive sentences (to the same extent as AA) shows that passive sentences are not only a consequence of reduced lexical competence, but also suggests that they are the result of pondering prototypical thematic roles, a process that takes place before encoding the lexical items and assigning grammatical relations. These results, then, lend credence to the idea of relational processes taking the lead in the planning of head-final RCs.

access, animacy and grammatical function assignment during this stage. In either case, both explanations are not mutually exclusive, and this stage can reflect a complex interaction: participants might have difficulties not only due to the retrieval of verb morphology, but also because both agent and patient must be kept active during verb retrieval and grammatical role assignment. The latter could be hampered either by lexical items with the same animacy status (AA condition), or by items that play the opposite roles they usually do (IA condition).

Note that, despite this difference, speech onset latencies between Japanese speakers and bilinguals do not differ. However, the underlying cause might be different: Japanese speakers start speaking on average after they have started preparing the second element. These participants seem to opt for the strategy of increasing accuracy over fluency, by preparing the whole sentence beforehand. This is not the case for bilingual speakers, who start their utterance after the first element has been prepared, and before the second is fixated. Thus, their lingering (and relatively longer times) would be likely more related to verb retrieval / lexical competition problems.

Finally, and also related to this temporal period, we analyzed the effects of voice on gaze pattern. To do so, we compared gaze patterns when bilingual speakers were going to utter (or were already uttering) active vs. passive sentences in sentences with the patient as HN. Once again, similar to the pattern found in monolingual speakers, results showed a tendency for speakers to look more extensively at the agent when the sentence was going to be an active than when the planned structure was a passive: that is, the agent was more focused when it was going to be the (subordinate) subject and the patient was more focused when it was going to be the subject of a passive sentence. This pattern was found at two different temporal points: when retrieving the first placed element, and after speech onset, when retrieving the second placed element. In other words, effects of subject planning were found *at the same time* lexical access was taking place. Exactly the same pattern was found for monolingual speakers: a pattern in which a general scaffold was created for the sentence beforehand, but was not fully completed until lexical retrieval of the corresponding elements was in order. It is surprising, once again, that this pattern was replicated with bilingual speakers whose native language has a very different construction. This result, thus, goes in the same direction of the pattern described in point 2 above: bilingual speakers are able to adapt to the planning requirements of their L2, and this can be accomplished thanks to a highly flexible production system.

In conclusion, this study has shown that bilingual speakers are able to adjust to the planning strategies that their L2 asks for, thanks to a highly flexible production system (Norcliffe & Konopka, 2015; Sauppe *et al.,* 2013; Konopka & Meyer, 2014). Speakers did not take longer than monolinguals to create a tentative structure, nor in starting to access the first uttered element. However, planning in their L2 was not cost-free, as they faced problems in accessing lexical elements (through competition between lemmas) that led them to devote extensive planning resources to both elements equally (reflected as equal proportion of gazes to agent and patient).

Chapter 6. General discussion

In this dissertation we have presented a series of studies exploring the planning processes that take place when producing relative clauses. We have compared the production of RCs by Spanish speakers and by Japanese speakers, and thereafter we have examined the planning processes shown by Spanish native speakers who had Japanese as a second language.

Spanish RCs resemble those of English: as a head-initial language, the Head Noun, that is, the syntactically most prominent constituent of the clause, appears in the first place, followed by the subordinate constituents. However, in Japanese, a head-final language, the opposite holds true: the HN appears at the end of the clause, despite its syntactic prominence in the clause, which is equal to Spanish. For that reason, these constructions offered an ideal framework to explore the different stages involved in sentence planning, while overcoming one of the main shortcomings of research in sentence production, namely, the fact that the first constituent of the sentence that is under analysis happens to be the most prominent (commonly the subject).

Language planning takes place incrementally. However, the precise nature of this incrementality is still under debate: linear incrementality accounts posit that lexical items are central in language planning. According to these accounts, structural relations are established after lexical elements are selected in order. On the other hand, hierarchical incrementality accounts submit that planning starts with a structural scaffold that guides subsequent planning and the retrieval of the corresponding lexical items. In this dissertation we were particularly interested in addressing several questions concerning the time-course of language planning that arise from this debate. First of all, in order to evaluate the tenets of these two accounts of incremental planning it is important to understand how the conceptual message undergoes linguistic encoding, that is, how speakers determine the “starting point” of the utterance and move from a non-linguistic message to the beginning of its linguistic encoding. Secondly, once linguistic encoding has started, it is necessary to clarify which information is prioritized when undertaking grammatical encoding. Finally, it is also important to elucidate how the processes of grammatical function assignment and constituent assembly involved in planning are related to the information retrieved from the mental lexicon.

In order to explore these issues, Japanese and Spanish native speakers were compared in our study during the time-course of planning, from its inception at the conceptual level to the articulation of the utterance. We did so by controlling the animacy of the elements involved (i.e. conceptual saliency), the type of RC (either with the agent or with the patient as HN), and by taking into account the voice of the output utterance (reflecting the assignment of grammatical roles). We explored three main points in time that correspond to the questions listed above: (1) The apprehension stage, where speakers decide the “starting point” of the utterance. It is at this point that the role of animacy in the initial comprehension of the scene is disclosed; (2) The beginning of grammatical encoding processes, where we should observe which information is prioritized, either structural relations (i.e. encoding from the syntactically most dominant element) or linear ordering (i.e. encoding from the first uttered element). (3) A later stage of grammatical encoding, where a relation is established between the assignment of grammatical roles to the sentence constituents and the retrieval of lexical elements in order, and where we purported to test whether there is an influence of grammatical role assignment in the planning of the lexical elements after lexical retrieval has started.

Moreover, in this dissertation, we explored the way in which bilingual speakers plan their speech in an L2, taking into account again these three stages, so as to ensure a better understanding of its relation with monolingual speakers planning. There are many remaining questions in bilingual sentence production, but our main aim was to analyze to what extent bilingual speakers rely on their native language, or are able to adapt to the planning strategies of the L2; or even whether they exhibit characteristics that are idiosyncratic of them as L2 speakers.

In order to explore these questions, we used eye-tracking measures, under the so-called ‘visual-world paradigm’. Gaze and speech have shown a tight relation (Griffin, 2004), with speakers focusing on the items they are looking at right before producing the words denoting them. It is a fine-grained method that enables to explore the time-course of planning from the conceptual arrangement of the message (through a provided picture to be described) to the moment that the utterance is finally produced. In this study, we exploited this relation to understand the planning of RCs, thus making this study the first attempt (to our knowledge) to explore the time-course of planning head-initial and head-final complex constructions from a cross-linguistic perspective, and also the first attempt to analyze how bilingual speakers of distant languages are able to overcome cross-linguistic differences in the planning of L2 sentences.

In what follows, we will present the results of our three experiments along with their implications and limitations in order, first by focusing on the behavioral data and then turning to eye-tracking data for the three different processing stages addressed in our study.

Analysis of the uttered sentences

Although in this study we included both RCs with the agent and with the patient as HN, prior studies that have explored the production preferences in RCs depending on animacy have always focused on RCs with the patient as the HN. This is due to the fact that these are the only ones that allow for voice contrasts, revealing the effects caused by the animacy of the elements involved.

RCs with the agent as HN consistently showed a high level of correct responses in bilingual speakers and a systematic use of active RCs in both monolingual and bilingual speakers, regardless of language and animacy combination. However, RCs with the patient as HN showed more variability. This type of RCs may be uttered in active or passive voice, with changes in the grammatical functions of the elements, but no change in word order whatsoever, as can be seen in (1) and (2).

(1)

a. Spanish object RC – active verb and patient as HN:

   La niña que empuja el niño lleva un vestido rosa

   The girl that pushes the boy wears a dress pink

   ‘The girl who the boy pushes is wearing a pink dress’

b. Spanish subject RC – passive verb and patient as HN:

   La niña que es empujada por el niño lleva un vestido rosa

   The girl that is pushed by the boy wears a dress pink

‘The girl who is pushed by the boy is wearing a pink dress’

(2)

a. Japanese object RC – active verb and patient as HN:

男の子が押している女の子はピンクのドレスを着ています。

Otoko-no-ko-ga oshiteiru onna-no-ko-wa pinku-no doresu-o kiteimasu

Boy-NOM is pushing girl-TOP pink-GEN dress-ACC is wearing.

‘The girl who the boy pushes is wearing a pink dress’

b. Japanese subject RC – passive verb and patient as HN:

男の子に押されている女の子はピンクのドレスを着ています。

Otoko-no-ko-ni osareteiru onna-no-ko-wa pinku-no doresu-o kiteimasu

Boy-DAT is being pushed girl-TOP pink-GEN dress-ACC is wearing.

‘The girl who is pushed by the boy is wearing a pink dress’

Despite the lack of changes in word order, both in Spanish (Gennari *et al*., 2012) and in Japanese (Montag & MacDonald, 2009) it has been observed an influence of the animacy of the patient: there is a tendency to assign the subject function to the animate patient (thus producing passive sentences), in contrast to cases where the patient is inanimate. In our results, we observed the same pattern. Analyzing the voice of responses depending on the animacy of the elements, it was observed that both in Spanish and Japanese, animate items were promoted to the subject position in RCs with the patient as HN, resulting in a higher proportion of passive RCs. Overall, there was a higher proportion of passive sentences in Japanese, when compared to Spanish, probably due to a general linguistic preference for passive structures in this language relative to Spanish, and to the lack of other linguistic choices that promote animate elements to a more salient position, as is the case of Spanish impersonal sentences (Gennari *et al*., 2012). Inanimate-Animate sentences (a combination that was not used in the Gennari *et al*. and the Montag & MacDonald studies) did result as well in an increased number of passive and/or impersonal sentences. Gennari *et al*. explained their results in terms of conceptual competition: AA sentences give rise to a higher number of passive sentences than AI because agent and patient are conceptually too similar, so speakers choose to separate them, either spatially or structurally. In this sense, the activation of the two nouns will take place early, and by assigning the subject function to the first one and the object (i.e. by-phrase) function to the second one, working memory load (needed for buffering these elements) is substantially reduced. Gennari *et al*. only analyzed head-initial languages, but the idea of competition could also be applied to Japanese, a language in which speakers would choose to assign the animate HN to the subject function despite its final position, or precisely *because* of its final position, in order to aid working memory and help the mapping of the different elements without interferences. However, our results with sentences with inanimate agents and animate patients (IA condition) cannot be explained by conceptual competition, as these authors propose, simply because the two elements involved in the scene differ to the same extent as the items in the AI condition do. This is not to mean that there is no conceptual competition involved in Gennari *et al*.'s results and perhaps even in our results in the AA condition. It is important to bear in mind that our research was not intended to analyze the effects of conceptually similar elements on language production; it was rather concerned with the effects of conceptual saliency on sentence planning. Hence, our behavioral results most likely point to a combined effect of conceptual competition in the AA condition (by means of promoting the HN to the subject position, and thus easing competition resulting from overlapping activation), and semantic prototypicality in the IA condition (by means of promoting the animate element, the HN, to the subject function, thus creating a more prototypical structure that, once again, will ease the planning of the utterance).

Note that it is also possible that in the case of IA sentences, speakers are simply promoting the animate element because it is the most salient one, without considering the whole conceptual structure, and opting for a more molecular planning by single lexical items, even if they are not uttered until the end of the clause, as is the case in Japanese. The only way to explore this is by analyzing the online planning process. We did so by recording speakers' gaze patterns as they planned their speech.

However, before we turn to the main results found in RC planning time-course, we will summarize the main findings observed in bilingual RC production in Japanese. Following my own Master's Thesis (Rodrigo, 2013), we wanted to explore the ways in which bilingual speakers of typologically different languages made use of conceptual saliency as a cue for speech planning. In 2013, we analyzed the production of relative clauses in Japanese by Spanish native speakers, in a written and an oral task. Our results showed that advanced learners showed a pattern that resembles that of Japanese monolinguals. However, an analysis of errors showed that most errors were related to the animacy of the elements involved: a tendency to assign the nominative particle to the animate element and the accusative to the inanimate element, regardless of the voice of the verb and their function as agent or patient, thus suggesting that animacy had an important role in planning RCs in this group of speakers. In the current research, we expected to replicate this pattern in the pattern of uttered responses. Our results were actually very similar to those: our bilingual group showed a higher proportion of passive sentences in AA and IA sentences than in AI sentences, but this proportion was closer to the one found with Japanese monolinguals than to that of Spanish monolinguals. Errors were, once again, related to the animacy of the elements: not only animate elements were assigned the nominative particle and inanimate the accusative particle more often, but also animate elements that were the patient were assigned the dative particle quite often. This is a case of transfer from Spanish (since direct objects that are animate are marked with preposition "a" (to) that roughly translates as the Japanese dative "ni"), but it also shows a tendency to separate the agent and patient roles soon and swiftly in order to avoid competition. In this dissertation, we wanted to explore whether this pattern, which we already observed in our 2013 work, would have any consequences in the way bilingual speakers plan their speech online.

Thus, subsequently, we analyzed the pattern of gaze movements directed to the agent and the patient from picture onset until articulation finished, for these three groups of participants. We will now discuss the results of this analysis in order.

Gaze patterns: online RC production

1. *Apprehension period: how is the starting point chosen?*

How the starting point of an utterance is chosen has been a matter of lengthy discussion. The importance of this step resides in the fact that the choices made at this point will likely determine the rest of the utterance, as Bock *et al.* (2004) point out.

Gleitman *et al.* (2007), Tomlin (1997), and Myachikov & Tomlin (2008), among others, have pointed out that visual cues that guide the attention of speakers to one of the elements result in more utterances starting with that element (as the sentence subject). That is, speakers start speaking with the first available element, without a prior apprehension of the whole scene. The structure of the utterance is constructed afterwards, as a result of the access to each of the elements in order. If this were the case, we would have expected that the animacy effect we observed in produced utterances should be reflected in early gazes to animate elements, in correspondence with an increased tendency to assign the subject function to those elements (thus producing passive sentences). However, despite the effects that animacy had on participants' responses, gazes patterns suggest that the conceptual saliency of the isolated elements (i.e. animacy) seems to play no role in the apprehension of the scene, as witnessed by the lack of effects on gaze patterns before 350 ms. following picture onset. However, when analyzing gaze patterns in RCs with the patient as HN, we observed a main effect of animacy on gazes to the patient in Spanish and a main effect of animacy on gazes to the agent in Japanese. In either case, pictures in the IA condition gathered a larger proportion of gazes, but did so to both agent and patient alike, and in both languages. On the other hand, sentences that were produced as passives gathered more gazes in general, to both agent and patient, and regardless of animacy. Thus, we did not find any interaction between animacy and voice: the likelihood of producing a passive sentence did not correlate with an increased proportion of gazes to the animate element vs. the inanimate element.

In the case of bilingual speakers, results slightly differed. They showed larger effects of animacy during the first time window: there was indeed an interaction between animacy and voice, although this interaction is difficult to interpret. Participants did indeed look at the animate element (the patient) more on IA pictures when they produced passives, but they also looked more at the inanimate element in AI pictures (the patient again) again when they produced passive sentences. The pattern responsible for the interaction was the one found in AA pictures. Surprisingly, in this condition bilingual speakers looked at the patient more often when they were to produce active sentences than the opposite, that is, when the patient (albeit the HN, which was held constant in the three animacy combinations) was not the subject of the sentence but remained as the object.

As we presented in previous chapters, bilingual speakers rely more strongly on animacy (Bates & MacWhinney, 1982; Kanno, 2007; Shirai & Ozeki, 2007, for Japanese) as a general cue to aid planning in a language that is not completely automated. For that reason, it was reasonable to expect that the animacy of individual elements would be a more central feature for bilingual speakers than for monolingual speakers, even from the very beginning, as a means to grasp the gist of the scene they had to describe. However, although participants acted differently across the three animacy combinations, it looks like the pattern differs from that predicted by strong linear incrementality accounts. It is difficult to have a clear understanding of the meaning of this gaze pattern. However, it might be argued that in this case, participants were influenced by the animacy of both elements, and when it was potentially confounding (AA pictures had two very similar elements), gazes were more concentrated in both agent and patient with a tendency to start looking earlier at the patient when it is not going to be also the subject. This might suggest that bilingual speakers were making use of a strategy of planning the HN earlier, in order to minimize potential competition.

To summarize, our results point more towards a brief apprehension of the whole scene, in which speakers understand the scene and identify the element they are being asked about in a very short period of time. In all three groups, gazes are shifted to the HN at around 300-350 ms., with slight differences between languages and groups with respect to the exact timing. Based on these results, it appears that a general gist of the scene, and of the structure to be produced, is accomplished very early in production planning. In the case of monolingual speakers, the fact that speakers of both languages were focusing more on the agent and the patient of the IA scenes (when compared with the other two animacy conditions) indicates that speakers are trying to understand the whole scene and its prototypicality. Animate elements are, then, selected as subjects not on account of their visual saliency, but rather because of their conceptual prototypicality, as was discussed with behavioral data. These findings are in accordance with Hierarchical Incrementality accounts. Not only eye-tracking studies, but also research on planning scope, show that, to use Bock *et al.*'s (2004) words, "speakers don't like to start what they can't finish". For example, Griffin (2003) showed that speakers were indeed able to start speaking after planning just one word, but this was only true when the length ensured enough time to plan the next word, that is, they did not start something that would not impede fluent speech. Similarly, Ferreira & Swets (2002) showed that speakers do not start counting the tens of a number without first ensuring they will be able to produce the following digits of the number. These results, along with ours with more complex structures, show that speakers are not oblivious to the whole conceptual representation of the message they want to express, or at least part of it (as shown by evidence that speakers can add contextual information afterwards (Brown-Schmidt & Konopka, 2015)). They do not start with the most salient element, without ensuring they can successfully continue the sentence. These results indicate that speakers analyze the coherence of the whole scene, pausing for a longer time on those scenes that are less prototypical or likely problematic (IA pictures) before shifting to the HN and start grammatical encoding. Even bilingual speakers, who seem to have a more complex relation with animacy from picture onset, do not rely completely on the animacy of each of the elements in isolation to decide their speech, but they scan the whole sentence focusing on elements that may cause problems, like competition in lexical access, as might be the case with AA pictures (remember that bilingual speakers showed longer time intervals devoted to lexical retrieval than monolinguals).

Note, finally, that participants in our study had to produce relative clauses, structures that allow little (if any) word order variation. It remains unknown what would have been the effect of the saliency individual elements on speech planning given complete freedom in choosing the starting word. However, what these results seem to suggest is that sentence planning starts with a global understanding of the scene that makes participants select structural relations according to prototypical thematic roles and the relation between elements.

2. *From apprehension to lexical retrieval: prioritized information in linguistic encoding*

After the initial 350 ms., speakers start the grammatical encoding of the utterance they are requested to produce. Whether this grammatical encoding is based on the retrieval of lexical items (*non-relational elements* –Konopka & Meyer, 2014–) or guided by a prior structural scaffold (*relational elements* –Konopka & Meyer, 2014–) is also not entirely clear. Once again, linear incrementality accounts and hierarchical incrementality accounts make different predictions concerning the way speakers undertake this step and which information is prioritized in it.

As stated above, RCs exhibit the opposite word order in Spanish and Japanese, the latter having a structure in which the HN is postponed to its subordinate constituents within the relative clause. This structure allows to analyze which information is prioritized in sentence planning, whether relational, and thus focused on the HN, as Hierarchical Incrementality accounts predict, or non-relational, and hence guided by the access to the first lexical item available, as Linear Incrementality accounts predict. If non-relational information is prioritized, we expect to find differences between both groups of monolinguals, due to opposite word order. In contrast, if relational information takes the lead, we expect that speech planning following early apprehension will be the same in both languages, since the structure to be prepared is the same.

Results showed that, despite cross-linguistic differences between Spanish and Japanese, the gaze patterns of monolinguals shared an identical form in both languages from 350 to 1000 ms. During this time window, gazes were directed to the HN, either the agent or the patient, depending on the type of RC. Notoriously, in none of the groups was there an effect of animacy on gaze patterns during this time window. This pattern of gazes in Japanese reflects the prevalence of structural over linear planning. In this language, the gaze pattern is reversed with respect to the order in which elements are going to be produced. After 1000 ms., gazes shift from the HN to the first uttered element, and then back again to the HN after 2500 ms., which is indicative of a relatively brief period of time devoted to retrieving the lexical item that will be placed in the first position. The structural representation that participants assemble in these time-windows guides subsequent gazes to both elements in the scene, so as to retrieve the corresponding lexical items in the appropriate word order. Thus, in Japanese, due to its head final nature, the preparation of the structure must precede the retrieval of the lexical items in order, showing a wide scope in RC planning (see Van de Velde & Meyer, 2014; Wagner *et al.*, 2010, Lee *et al.*, 2013, for evidence of scope planning in complex structures). On the other hand, in Spanish, participants show the same initial pattern as Japanese speakers, with speakers focusing first on the HN. However, in Spanish the HN happens to be also the first placed element. With this data alone it is not possible to tell whether Spanish participants are creating a structural representation of the sentences before retrieving the corresponding lexical items, as seems to be the case in Japanese. Lee *et al.* (2013) showed that RCs are planned together also in head-initial languages. In their study, they controlled the codability of the elements involved in the subordinate clause to measure the ease of speech onset. We did not control for word frequency, as it was not the aim of our study (although we ensured that words were known to L2 speakers). However, we presented all the nouns and verbs before the experiment, and the verb was shown again right before each picture. This likely resulted in easier access to lexical items and to the structure (see Konopka, 2012; Ganushchak *et al.*, 2014 and Van de Velde *et al.*, 2014, for the effects of previously seen nouns and verbs on speech planning). Nevertheless, this does not mean that structural planning was necessarily simplified or removed. In any case, it seems likely that lexical retrieval started earlier in Spanish than in Japanese, in a more interwoven fashion, due to the simple fact that the Spanish grammar allows it. Thus, when there is no restriction or hindrance, speakers can make use of all the available information to be able to plan their utterance efficiently. This difference reflects the considerable degree of flexibility in language planning mechanisms (e.g. Norcliffe & Konopka, 2015), with speakers efficiently focusing their gazes to the elements they have to prepare in accordance with the structural requirements of each language ("seeing for saying", as Bock *et al.*, 2003, note).

Bilingual speakers, in turn, showed a very similar pattern to that of Japanese monolinguals. They started focusing the HN at around 350 ms. and continued to do so until 1000 ms., before shifting to the first mentioned element. The high accessibility of the structure mentioned above (due to the verb being presented right before the picture) might have helped bilinguals to swiftly access the HN after apprehension. Konopka & Forest (2016) found the same pattern. They observed that bilingual speakers take longer to start lexical encoding when producing transitive sentences. However, this delay disappeared when participants had previewed the verb they had to produce, making the structure more readily available. In contrast, the preview of the first noun did not affect production timing. Our results replicate these findings, showing that structure planning takes place easily, while lexical retrieval seems to be the main source of problems for bilinguals.

These results also show that bilingual speakers can adjust their planning strategies to those required by their L2, even in a grammar completely different from their own L1. Once again, we can conclude that speech planning is flexible, not only between speakers, but even within the same speaker in her different languages.

One point of concern about the evidence of preferential looks to the HN found in both languages is that it may reflect a bias introduced by the question provided to our participants, which asked about a particular feature of the item denoted by the HN. It might thus be argued that participants were just tracing the element they were being asked about. However, this possibility seems unlikely, as it has been shown that speakers can locate the first information relevant to prepare the utterance (Bock *et al*., 2003), or the patient of an action (Griffin & Bock, 2000), in a period as short as 300 ms. Hence it looks highly improbable that our participants took as much as one second just to locate the item they were being asked about before getting involved in linguistic encoding processes[48]. Additionally, the possibility that initial gazes to the HN were the result of a conscious strategy in Japanese speakers, who take longer than their Spanish counterparts in starting to speak, also seems unlikely. In Study 2, with Japanese speakers, we presented an additional analysis including only speakers whose speech onset was shorter than 3 seconds. These Japanese "fast" speakers showed exactly the same pattern as the one found when including both "slow" and "fast" participants. If any, the most remarkable difference was that "fast" speakers started speaking as soon as they prepared the first element they were going to utter (i.e. the subordinate NP), showing a more incremental planning process. These speakers are more similar to Spanish monolinguals in that respect.

In conclusion, it seems that, at least under certain circumstances, structural information is prioritized when undertaking linguistic encoding in RC planning from the message's conceptual representation, in both monolingual and bilingual speakers. These results point to a rather flexible system that allows the creation of the structure of the whole clause before retrieving the corresponding lexical items (i.e. a wider planning scope), but also allows preparation of the lexical items that are going to be produced at the same time (or at least, overlapping in time to a certain extent) as the

[48] Ganushchak *et al.* (2014) also conducted an experiment in which they focused their participants' attention by providing a question: e.g. "What is the policeman stopping?" Although their results show some differences with respect to ours, they also indicate that the search of the object that were being asked about took place fast and efficiently. One of the main differences between their results in Dutch (object-focus, which is the most similar condition to our RCs in Japanese) and our results is that in their case, participants barely looked at the subject once they created the structural scaffold. This is so because the lexical item for the subject ("the policeman") was provided in the question, so that name-related gazes were not needed. Note that we did not include the name label of any of the elements in the questions.

creation of the structure in hand is taking place. Whether or not grammatical functions are completely assigned before linearization begins is the target of our next and final issue.

3. *Speech planning during lexical retrieval: is there an interaction between Grammatical Function Assignment and Constituent Assembly at this stage?*

The last comparison we carried out in this study was focused on the relation between grammatical function assignment and constituent assembly, once lexical retrieval has got started. With that purpose, we focused on RCs with the patient as HN, and compared active and passive responses in each language, looking for differences related to voice selection, which entails the assignment of grammatical functions to each constituent. In this case, the comparison we made was within languages. In both languages, word order is identical between active and passive sentences in RCs with the patient as HN. Our main concern was to establish whether grammatical function assignment could be teased apart from constituent assembly, and whether they overlapped in time. Results in the previous point showed that the structural scaffold (Ferreira & Bock, 2014) was constructed before lexical retrieval started, at least for Japanese monolinguals and for bilingual speakers. However, these results did not show how precise this scaffold is. Lee *et al.* (2013) showed that the ease of access to lexical items within an RC that are located far away from each other affects RC planning, thus suggesting that lexical retrieval of all the items has at least partially started. With lexical retrieval comes the assignment of grammatical roles to each constituent.

In this regard, our results point towards a kind of planning in which the assignment of grammatical roles still lingers as the retrieval of lexical items has begun (i.e. as participants focus on the element they are going to produce first). In both Spanish and Japanese, and noteworthy, in bilingual speakers in their L2 Japanese as well, voice selection correlates with differences in gaze patterns between active and passive sentences. However, the exact timing in which these differences arise differs between speakers.

From 1000 to 1800 ms. there is an interaction between animacy and voice in Spanish, as animacy modulates the exact timing in which grammatical functions are assigned. These differences remain from 1800 to 2500 ms. as shown by an interaction (for gazes to the agent) and a main effect of voice (for gazes to the patient). On the other hand, differences between active and passive sentences in Japanese appear from 2500 ms., and persist beyond SO, reaching as far as 6000 ms. Despite the differences found between both languages in timing, duration and strength of voice effects, in both cases differences between active and passive sentences are visible around two seconds before SO, and show the same gaze patterns. In both languages, these effects emerge when participants are preparing the lexical item of the HN (the patient), which is the first mentioned in Spanish, but the second one in Japanese.

In general, the main pattern in monolingual speakers reflects a convergence of gazes to both agent and patient when an active sentence is going to be produced, in contrast with passive sentences.

In the case of bilingual speakers, they also show differences in gaze patterns due to voice, but these differences do not show up until later in the planning process: there are main effects of voice from 4800 ms. onwards, with no interaction with animacy. Thus, in this case, differences due to voice assignment (or grammatical role assigment) do not start until speech has begun. However, in accordance with the patterns found in monolingual speakers, voice effects are present from the moment speakers start lexical retrieval of the HN, a process that starts later for bilingual speakers than for Japanese monolinguals. Moreover, the pattern of gazes is exactly the same as that found in monolinguals: the proportion of gazes to agent and patient are balanced when speakers are starting to produce an active sentence, while there is more difference between the looks to agent and patient when a passive is going to be uttered. In other words, agents receive a greater proportion of gazes when the sentence to be produced is an active that when it is a passive. Note that this pattern occurs despite the agent not being the main focus at that moment, as speakers are involved in retrieving the lexical item for the patient.

Different reasons can underlie this pattern. In the first place, this pattern of convergence of gazes to both elements may reflect greater difficulty in integrating structural information in sentence planning (Norcliffe *et al.*, 2015). In this case, in active RCs with the patient as HN, speakers are seen to have more trouble integrating the subordinate agent as the subject of the sentence. This greater difficulty could also explain the low proportion of active sentences when the patient was the HN, both in Spanish and in Japanese for both L1 and L2 speakers. This greater difficulty is also supported by speech onset latencies in Japanese L1 and L2 for relative clauses with the patient as HN. When RCs are produced in active form, speech onsets are longer than when they are produced in passive form, showing more complexity and planning difficulty. However, this difference in SO latency does not occur in Spanish, a fact that is not easily explained if we assume that planning active sentences is more difficult across the board. Word order could have an impact here, since in Japanese voice planning begins after the first element has been prepared and even after speech onset, while in Spanish, this difference shows up in the planning of the first element, which gives time to recover from any difficulties in planning active RCs with the patient as HN However, the precise way in which this could have influenced our results remains to be explored.[49]

Another possibility, although less plausible, is that the convergence of gazes is not due to planning difficulty, but simply to the fact that speakers focus on the agent in active sentences in order to assign this element the subject function when it is incorporated to the ongoing scaffold. Note, however, that the assignment of the subject function could take place during the retrieval of the patient. In the case of Spanish, this results in the assignment of the grammatical role before the lexical item corresponding to the agent has been accessed (before gazes have shifted to the agent). However, in the case of Japanese (L1 and L2), assignment of the subject role to the agent should have taken place retroactively, that is, after the lexical item corresponding to the agent has been accessed. For that reason, this possibility seems unlikely, since it is precisely Japanese the language that requires earlier commitment to grammatical roles in nouns, due to its case particle system[50] (see Myachykov & Tomlin, 2008 for results in Russian −a case-marking language− showing an earlier assignment of grammatical functions).

Regardless of the reason that might underlie this pattern, it is remarkable that, despite the fact that word order is exactly the same between active and passive

---

[49] Note that, despite the differences between both experiments, results from Konopka & Kuchinsky (2015) seem to converge with ours. They found an increased proportion of fixations (sharper fixations) to the agent when speakers produced active sentences in a SVO order after active primes in the time window from 400 to 1500 ms., denoting that the accessibility of a structure led participants to direct more attention to the element under preparation (the agent) without having to devote more resources to structure building.

[50] Remember that Japanese have case particles (nominative, accusative, dative, etc.) that must be attached to each noun of the sentence and signal its grammatical function: topic, subject, direct object, oblique object.

sentences *within* languages (that is, the word order of Spanish active RCs is the same as Spanish passive RCs and the word order of Japanese active RCs is the same as Japanese passive RCs; see examples (1) and (2) above), there are differences in gaze patterns arising from differences in the grammatical function assigned to each of the items involved in the scene. Thus, even when a tentative scaffold of the structure is created at early stages of production planning, with structural relations guiding gazes to the different elements, it can be argued that grammatical relations are not fully established before lexical retrieval begins. This suggests a kind of model in which the processes of grammatical function assignment and constituent assembly overlap to a certain degree in a cascaded fashion process, in accordance with a weaker version of the Hierarchical Incrementality model.

In the next section, we will try to analyze these results globally, in order to elucidate how they can be incorporated to current models of sentence planning. In the final section of this chapter, we will make some comments on the limitations of our studies and will suggest possible directions for future research. In our discussion of the implications of our current studies, we will separately address the results of L1 and L2 production planning, not because we think there is an abrupt difference between them, but because models of planning have traditionally focused on just one of these types of planning. We will, nonetheless, try to integrate both as far as we can.

Implications of our results

<u>*L1 speech planning*</u>

The comparison held in this dissertation between planning strategies in a head initial and a head final language stands in the middle of the debate on whether speech planning is lexically or structurally mediated. There is consensus in that these two factors play a prominent role in both message encoding and speech planning. In Chapter 2, we presented a description of the language production system proposed by Ferreira & Slevc in 2007. In this model, both lexical access (left part of the scheme) and structural assembly (right part) are directly connected with message encoding (i.e. the non-linguistic representation of the utterance speakers want to express). However, which is more crucial or which is prioritized in this process remains unclear. Incrementality is assumed in both cases, although the type of incrementality differs in each case (Ferreira & Bock, 2014): linear incrementality, prioritizes lexical access, whilst hierarchical incrementality gives priority to structural relations between elements. The results in this study are more in line with a hierarchical incremental process. To begin with, these results show that conceptual accessibility of single lexical elements did not exert a direct influence on the assignment of grammatical relations, but did so in terms of the prototypicality of the whole scene. A short apprehension stage, which only lasted up to 350-400 ms. and where agent and patient were similarly focused, preceded a focus towards the HN. In Japanese, this meant that speakers were extensively fixating an element that was not going to be placed at the beginning but that was, nonetheless, the most prominent element in the scaffold they were starting to plan. In the case of Spanish, we face the same criticism that has been raised against other studies making use of transitive clauses, namely that it is not possible to differentiate whether participants are embarked in the assignment of structural relations exclusively or whether lexical access starts along with the planning of the HN. Cross-linguistic evidence (e.g. Sauppe *et al*., 2013; Norcliffe *et al*., 2015; Kubo, 2016; Ganushchak, *et al*., 2014; Norcliffe & Konopka, 2015) shows that there is flexibility in the way in which different languages handle the relation between both processes. Thus, it could be the case that in Spanish incrementality unfolds in a more linear fashion from the very beginning, simply due to the fact that its grammar allows it. However, what seems certain is that speakers started fixating the HN after a prior understanding of the scene, subsequently directing their gazes to the HN initial element.

It is important here to differentiate the effects of conceptually accessible elements in planning from the use of relational vs. non-relational information along the planning process. The former refers to the transition from the pre-linguistic message to the realm of linguistic encoding. In this regard, our results suggest that in both languages, this transition is not mediated by single elements that are more salient in the scene (at least, under our non-contextual, controlled experimental framework), but rather by the apprehension of the whole scene. Lee *et al.* (2013) already showed that planning of the whole message comprised within the relative clause takes place beforehand, with a wide planning scope in these structures. Our findings, especially with Japanese, support their results. This kind of message planning creates a structural scaffold, but this scaffold is tentative and open to unexpected changes (Ferreira & Bock, 2014).

After message planning has successfully given way to linguistic encoding processes, there starts the “waltz” between relational and non-relational information. Our results suggest that there is a rather clear-cut distinction between early scaffolding processes and lexical retrieval in the case of Japanese, thus adding a nuance to the presumed tight gaze-speech relation (e.g. Griffin, 2004; Konopka & Brown-Schmidt, 2014): there is an anticipated encoding of the structure, with speakers focusing on the most dominant element of the structure, followed by a retrieval of the lexical items in order, an order that reverses syntactic dominance relations. In other words, our results show that there can be an extensive focus on one of the elements that is syntactically dominant, yet not the first to be mentioned, before name-related gazes start. However, in the case of Spanish, a successful early structural scaffold can “call” the corresponding lexical element without waiting until 1000 ms., as in Japanese, opting for a more interwoven way of incremental planning.

However, this is not to mean that in Japanese all structural assembly takes place at the very beginning and all lexical access takes place later, since this would violate incremental planning assumptions and result in non-fluent speech. Our results regarding voice differences further endorse this. In both Spanish and Japanese, there were differences due to structural assembly once lexical retrieval has started, specifically when the lemma of the HN (the patient, in this case) was being retrieved. Active RCs with the patient as HN were more difficult to assemble, showing greater convergence of gazes between agent and patient at this stage. What this suggests is that early structural assembly is not fully complete and serves as an anchor for the retrieval of lexical items. This incompleteness is, precisely, what allows its flexibility. As the planning process evolves, relational and non-relational information unfold in an interwoven fashion in both a head initial (Spanish) and a head final (Japanese) language. Figure 6.1 shows a representation of the process we have proposed.

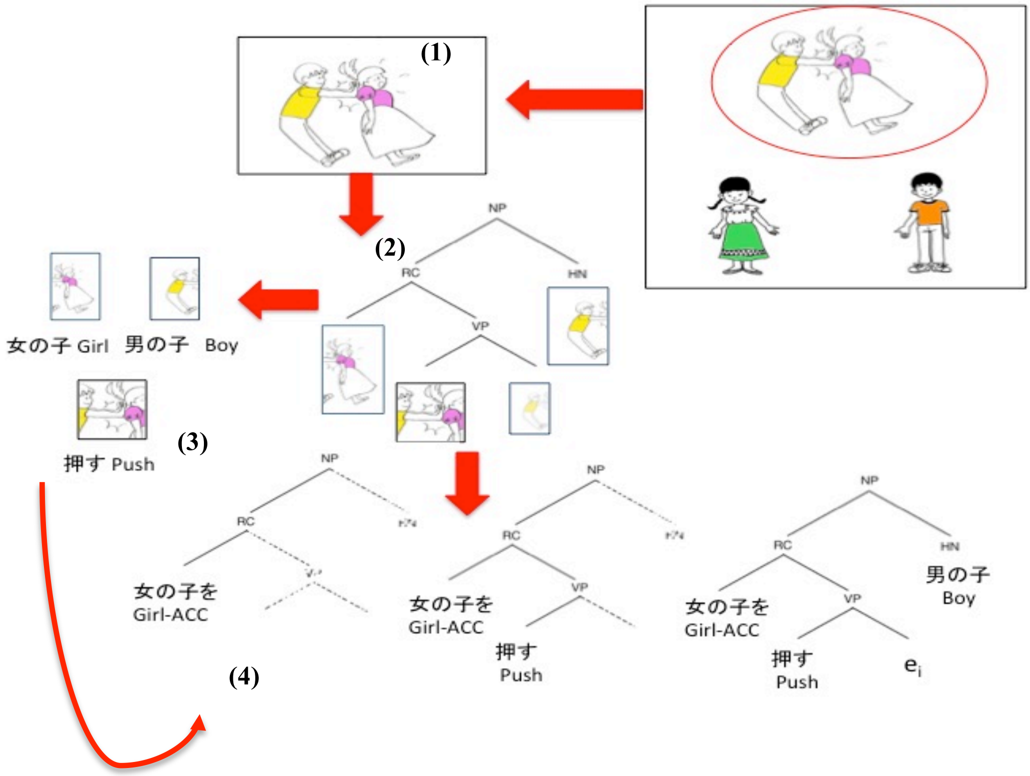


Figure 6.1. Representation of the process that is supposed to occur in Japanese: (1) there is an apprehension stage in which speakers experience a global understanding of the whole scene and create a message to utter. After that, in the second stage we defined (2), the message undergoes a structural scaffold that subsequently (stage three) (3) incrementally retrieves the corresponding lemmas from the lexicon, and assembles them (4) in an structure in order with the corresponding selected grammatical functions (in this example, the object function for the noun "girl", marked with accusative case, and the active form for the verb "push").

Finally, our results also provide further evidence for animacy effects on RC planning. We added an animacy combination not included in previous studies on conceptual accessibility: inanimate agent – animate patient. This combination yielded the same amount of passive RCs with the patient as HN as animate-animate scenes did, both higher than animate-inanimate ones. Our results replicate those by Montag & Macdonald (2009) with Japanese, and Gennari *et al.* (2012) with Spanish, and further provide evidence that not only the similarity between agent and patient (in the form of conceptual competition, as Gennari *et al.* note), but also the general prototypicality of the whole scene –with inanimate elements being less likely to take the subject function– have effects at this late stage of planning. Eye-tracking results showed animacy effects in interaction with voice choice, showing that the likelihood of a given item of becoming a subject, as well as the similarity between participants affects structural planning while lexical retrieval is taking place. Speakers show a greater convergence of gazes when voice planning is taking place, which shows once again that grammatical role assignment is not based solely on the animacy of one of the elements, but comes about as a result of contrasting both elements in terms of their resemblance and their most prototypical roles.

L2 speech planning

In this study we also analyzed the planning process of bilingual speakers, whose L1 was Spanish and L2 Japanese. The grammatical system of their two languages completely differs, and so does the word order of RCs, forcing native Spanish speakers to reverse the way in which they plan their speech, or at least the preparation of the lexical elements involved in it.

Despite the completely opposite word order, advanced bilinguals showed no problems in adapting to the requirements of the L2. Speech onset was larger than in Japanese, although it widely varied with type of RC and animacy combination, with some of the combinations showing a similar SO to that of Japanese natives. Speech patterns showed that bilingual speakers only had some problems in accessing the lexical items, but not in apprehending the scene (which was expected, as it is a language-independent process, e.g. Levelt, 1989), or in constructing a structural scaffold based on this apprehension focused on the HN prior to lexical access.

The fact that structural scaffold takes place efficiently also in bilingual speakers suggests that bilingual speakers are competent in connecting the conceptual information to structural relations prior to accessing specific lexical items. In our study, we provided the verb that participants had to use beforehand, thus facilitating the accessibility of the structure. Konopka and Forest (2016) showed that when the structure is accessible for bilingual speakers, planning starts without delay. Our results replicate this pattern. The results with bilinguals in Japanese show more strikingly the separation between the two processes of structure scaffolding and lexical retrieval. Lexical retrieval was overall problematic for bilingual speakers (maybe this shows that we did not train them properly), but this problem is not reflected in the first 1000 ms. of planning, but only after that, once bilingual speakers start fixating the elements in the scene. However, similarly to monolingual speakers, this separation does not mean that bilingual speakers decide to prepare the whole structure before searching the lexical items that fill that scaffold. This process could have resulted from a problematic planning that was too costly for bilinguals, leading them to follow the strategy of giving up incremental production in favor of efficiently producing the utterance they were asked for. The effects due to voice assignment present in bilinguals, however, indicate that this is not the case. Bilingual speakers showed the same differences due to voice assignment as monolingual speakers of both groups: active sentences were more complex and yielded longer speech onset latencies and more divided gazes between agent and patient than passives.

Perhaps the similarity between bilingual speakers and Japanese monolingual speakers is the most remarkable result from our study. Errors due to interference coming from the L1 were almost entirely lacking, and were mostly related to the animacy of the elements. This lack of interference between languages was also seen in Rodrigo (2013), but in this study we added the finding that planning process resembles that of Japanese, without any major disruptions coming form Spanish.

Lack of effects of the L1 on the L2 on RC constructions can be explained by, at least, two of the three models we presented in Chapter 3. De Bot's model poses that a bilingual's two languages possess two different formulators that are connected to each other. However, the more distant the two languages, the weaker the connections are claimed to be. Japanese and Spanish are two languages that share virtually no grammatical features, making the connections weak and not visible in the face of an experiment that explored the effects of conceptual information on language, rather than the connection between linguistic systems. Secondly Hartsuiker & Pickering's (2007) model can also explain the lack of interference. These authors posit that a bilingual's two languages are interconnected with the same strength, regardless of linguistic distance, given enough proficiency. However, there is one requirement that must hold for this connection to happen: the linguistic structures in both languages must be formulated in the same word order. When word order varies, as it is the case in our study, connections do not hold, and so interference vanishes. However, and contrary to Hartsuiker & Pickering's lexicalist approach, the fact that bilingual speakers' early representations seem to be focused on the structure rather than on lexical elements involved plays against their assumption of planning based on the activation of lexical nodes. Once again, we cannot be sure of which of these possibilities prevails, as we did not focus our efforts on the connection between structural representations between languages. In the next section, we will present some future directions to further explore this issue.

In conclusion, the main findings of our bilingual study are that bilingual speakers of typologically distant languages can shape their planning strategies to adapt them to the requirements of their second language. We claim to have shown that under certain circumstances, and as we presented in Figure 6.1, planning of the structural scaffold precedes lexical activation, but the structure is not completely created beforehand. A same speaker, in her own two different languages (even if acquired in adulthood) can modify the extent to which their utterance is structurally or linearly incremental from the beginning, which drives us a step farther in the understanding of cross-linguistic flexibility in language incremental production (e.g. Norcliffe & Konopka, 2015).

Limitations and future directions

The experiments presented in this dissertation suffer from some limitations and have left some open questions too, both in monolingual and bilingual studies. In this section, we will examine the main limitations and will explore possible ways to overcome them, along with some future directions for research, both in short-term and long-term studies.

In the first place, one limitation, which is inherent to research on Spanish RCs, is that HN planning cannot be differentiated from first-element planning. Our experimental setting was aimed to compare Japanese and Spanish, and moreover to compare monolingual with bilingual speakers. For that reason we used short RCs and did not control for the codability of the structures and the frequency of lexical items, but only for conceptual accessibility. Thus, our current results do not allow us to conclude that exactly the same pattern is shared between languages, nor can we know to what extent cross-linguistic differences can be traced. Further studies should focus on RCs with an approach similar to the one shown in Lee *et al.,* controlling for elements that are distant from each other in the sentence, but closely related structurally. By means of the visual-world paradigm, and comparing Japanese and Spanish, we could observe the size of planning scope in these two languages, and to what extent they resemble in their way of interweaving relational and non-relational information.

This leads us to a second point of concern on which we would like to focus in future studies. Japanese speakers showed a planning focused on the HN, which led to subsequent lexical access. In other words, scope planning was wide enough to encompass the whole conceptual relations involved in the relative clause, before starting access to lexical items. The upper limit of this “reversed order” planning shown by Japanese speakers and bilinguals remains unknown. In other words, what would happen with longer or more complex RCs (for example, double antecedent RCs)? Would planning still revolve around the HN in a language like Japanese?

More importantly, however, is another major concern we have held with this study. Relative clauses are, by nature, structures in which one of the elements is the focus, an intended referent about which the speaker wants to say something. In order to prompt these type of sentences, and based on studies that focus production of RCs (eg. Gennari & MacDonald, 2008; Montag & MacDonald, 2009; Gennari *et al.*, 2012), we asked the participants beforehand about one of the elements involved in the action, either the agent or the patient. This involves an obligatory search for the HN prior to start linguistic encoding, a search that could have caused, or at least, enhanced the pattern of gazes to the participant denoted by the HN prior to lexical retrieval, as it happened in Japanese monolinguals and bilingual speakers alike. In Chapter 4, when we discussed our results in Japanese, we submitted that we deemed this possibility unlikely, since previous research has found that referents can be detected in less than 400 ms. (as in patient-detection condition in Griffin & Bock (2000)’ study). However, in order to be completely sure that a search for referents is not behind these results, we are planning to undertake a follow-up study in the near future. In this study, we will ask participants to produce RCs without a prior question that may lead their gazes. An example can be seen in Figure 6.2.

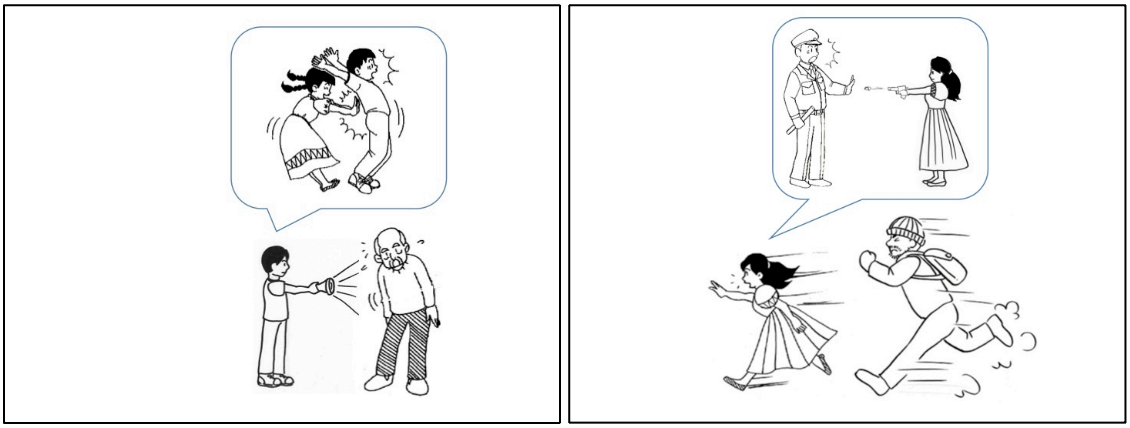


Figure 6.2. Possible pictures for the follow-up study. In the picture on the left, an object RC that modifies the subject is expected ("The boy who the girl pushed is lighting the old man"). In the picture on the right, a subject RC that modifies the object is expected ("The thief is chasing the woman who shot the policeman"). Passive sentences in both main and subordinate clause are possible and will be allowed as long as the content of the subordinate action (shown in the bubble above the main scene) is produced as an RC modifying the appropriate constituent of the main clause (signaled by a pointed angle).

Participants will have to describe images like those shown in Figure 6.2, stating what the boy (left) or the woman (right) had done before the time of utterance (as represented in the speech bubble), while describing the main scene. As we explained above, RCs are inherently focused and contrastive, so in order for participants to express the content of the speech bubble as an RC, they will be instructed to believe that another participant will later to pick out the correct picture from a choice of two alternative options by listening to their descriptions. They will be told that the main scenes will be held constant while the characters in the subordinate (past) actions will vary, so participants will have to be very explicit about which particular character was the doer or the recipient of the action (e.g. they will have to be clear about which boy is lighting the old man, or which woman is being chased by the thief). Practice trials showing actual contrasts will help us prepare the mindset for contrastive descriptions, without prompting and explicitly biasing participants' responses, nor having to show them the alternative pictures supposedly displayed before the other participant. An example of a practice trial is shown in Figure 6.3.

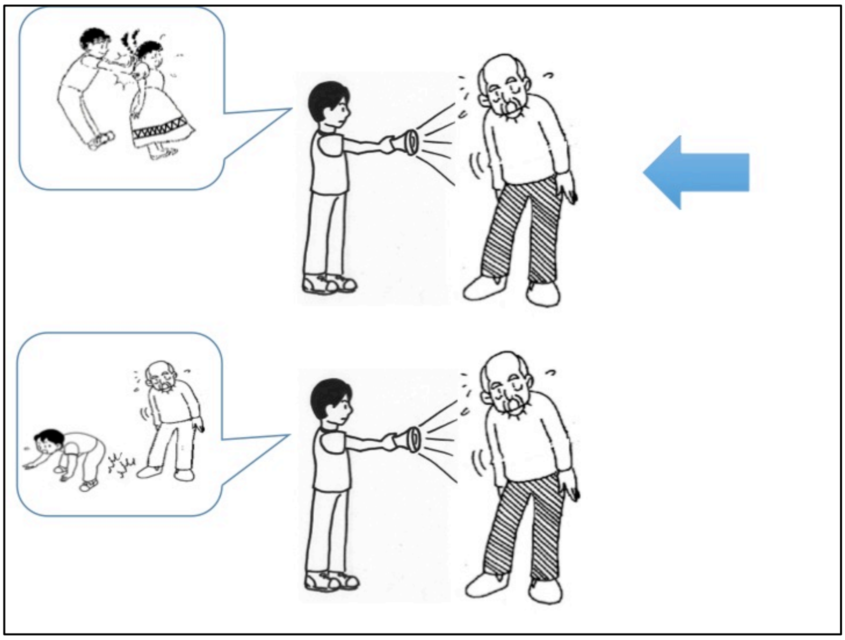


Figure 6.3. An example of picture that will be used during the explanation of the task and the practice trials. Participants will be taught to think that another participant (not present in the room) will see a picture similar to this one, with two choices but without the arrow, and will have to choose the correct picture being described. Participants will be told that the only difference is what the boy had done in a past event (depicted on the bubble), that is, which boy was it that lit the old man. They will be told that they will only see one of the two contrastive pictures during the task (the one with the arrow), but they must keep in mind that they will be describing the pictures for another participant to make his/her choices.

As can be seen in Figure 6.2, both RCs with the agent and with the patient as the HN modifying in turn the agent or the patient of the matrix clause will be included, resulting in a 2 × 2 design (type of RC –agent vs patient as HN– × modifying element of the main clause –agent vs patient–). The aims of this study are twofold:

1. To explore the planning process of RCs in Japanese and Spanish without a question and a verb provided before each picture, so as to examine planning processes in a more ecological setting. We will explore whether the process is the same or differs from the one found in the studies included in this dissertation.
2. To analyze the size of planning scope in Japanese, in contrast with Spanish, especially by analyzing RCs that modify the agent of the main clause with those that modify the patient. Analyzing full utterances, like the ones presented in (3) and (4) (which correspond to the pictures shown in Figure 6.2)

will help examine the process of incremental planning and the extent of the structural scaffold in RC planning (RC in brackets):

(3) Agent-modifier RC with patient as HN:

［女の子が押した／女の子に押された］男の子がおじいさんを照らしている。

[Onna-no-ko-ga oshita / Onna-no-ko-ni osareta] otoko-no-ko-ga ojiisan-o terashiteiru.

[Girl-NOM pushed / Girl-by was pushed] boy-NOM old man-ACC is lighing.

"The boy who the girl pushed / was pushed by the girl is lighting the man".

(4) Patient-modifier RC with agent as HN:

泥棒が［警察官を撃った］女性を追いかけている。

Dorobou-ga [keisatukan-o utta] josei-o oikaketeiru.

Thief-NOM [policeman-ACC shot] woman-ACC is chasing.

"The thief is chasing the woman who shot the policeman".

In this planned study, we will not control for the voice of the responses, since natural production will be prioritized. However, there is another question that may arise from the analysis of the results and that we consider to be interesting to explore in a near future. A problem present in voice effects is that we did not control for the voice of the responses. As a result, the proportion of active and passive RCs varies widely between animacy combinations and, to a lesser extent, between languages and participant groups. Forcing a voice response would hinder natural planning process, but would allow a more careful analysis of the structural planning that takes place after lexical retrieval has started: what is the exact timing in which these differences appear? Will voice effects still interact with animacy even in a forced-choice task? Would such effects be comparable between languages, and L1 and L2 groups? We expect that the main effects and timings will remain even with forced choice of voice, but interaction with animacy might vary with it. The nature of this interaction and the ensuing differences remain to be fully understood.

Regarding the study with bilingual speakers, the most obvious improvement would be to increase the number of valid participants. One of the major caveats of our study is the low number of valid participants and responses. This problem might have masked or distorted differences or trends, or even make small differences more salient. In general, we predict the main patterns will hold, but voice differences and marginal differences can show a different face that is important to explore.

However, the next step in our research with bilinguals does not only involve a mere increase in participants, but we also consider it important to explore the relation between L1 and L2 concerning this structure. Our results showed almost no influence from the L1, but this lack of influence deserves an explanation. Maybe there is simply no connection in the planning of RCs, as Pickering & Hartsuiker's model would predict, given that the word order of RC in both languages differ, or alternatively, it might be the case that since participants in our study already reached an advanced level, they were able to suppress or reduce the relation (or interference) between languages, and making it invisible to our study, which was not intended to exploit these connections. For that reason, we think a subsequent study should focus on the relation between languages, by means of a structural priming study paired with eye-tracking, so as to observe the effects not only on the uttered responses, but also on the online planning process. This way, we will not only be able to understand the planning process that is inherent to the L2, as we observed in this study, but we will also have a clearer idea of what the relation is with the L1 all along the process.

In this connection, regarding bilingual speakers, our current study yielded a static picture of the planning process of RCs by advanced bilinguals. However, it is equally important to analyze the development of planning in the interlanguage of bilingual speakers. In the transition from intermediate to advanced speakers, it makes sense to ask whether there is any change or evolution in the way speakers combine relational and non-relational information while planning in their L2. We are aware that RCs is a kind of construction that poses difficulties for our intermediate speakers to produce it smoothly and fluently enough to allow eye tracking measures. We are exploring other complex structures that will suit our purposes, yet manageable for intermediate-level learners.

Finally, another remarkable conclusion from these studies was the flexibility shown in speech planning in RCs, both between speakers of different native languages, and within the same participants when speaking a different language. An interesting follow-up would be to explore RC planning within a language that allows both HN-initial and HN-final relative clauses. For that purpose, we considered Tagalog to be an adequate language. The following examples, from Aldridge (2003), show two different word orders that RCs in Tagalog can show, namely head-initial (5) and head-final (6):

(5) *libro*-ng b-in-ili ni Maria

book-LK -PERF-buy ERG Maria

'the book Maria bought'

(6) b-in-ili ni Maria-ng *libro*

-PERF-buy ERG Maria-LK book

'the book Maria bought'

Prior studies with transitive clauses in Tagalog (Sauppe *et al*., 2013), Tzeltal (Norcliffe *et al*., 2015) and Kaqchikel (Kubo, 2016) showed differences in the way the same participants planned verb-initial and subject-initial sentences. We expect that a similar flexibility will arise here. Moreover, if the same structure-focused planning that was observed in Japanese is observed in Tagalog head-final RCs, we could extend these results to other typologically different languages.

Conclusions

The development of eye-tracking measures that allow recording of eye movements while talking has helped to acquire a deeper knowledge of the way in which sentence planning takes place along its time-course. This dissertation was intended to explore some of the unresolved issues regarding the information used and the timing in sentence planning by monolingual and bilingual speakers. Along the three studies presented we have examined the planning process of RCs first from a cross-linguistic perspective, comparing monolingual speakers of Japanese and Spanish, and then by exploring bilingual sentence production, with a group of Spanish-Japanese late bilinguals.

Our main aim was to compare two alternative accounts of sentence production planning, namely, hierarchical incrementality and linear incrementality, and analyze how relational (structural) and non-relational (lexical) information is used along the whole planning process in a HN-final and a HN-initial language. We hypothesized that a planning process that prioritizes structural relations will yield similar results in both languages, despite linear (word order) differences, while a planning based on lexical items will show cross-language differences from the very beginning. Moreover, we wanted to explore how bilinguals of these two languages showing the opposite word order plan their speech in their L2.

Several conclusions along the three studies are worth summarizing at this point. In the first place, when speakers are faced with the task of producing a sentence (either when answering a question, as it was our case, or for internal motivations), they create a prior non-linguistic message that captures a general idea of what they have to convey. This global encoding has been found to be oblivious to the animacy of single elements, but is somewhat sensitive to the distribution of the animacy feature in the whole scene, as this feature represents its prototypicality. There was almost no effect of animacy on gazes along the whole process, although animacy combinations did modulate the amount of active or passive responses in those RCs with the patient as HN, suggesting once again that it was not the accessibility of a single lexical element, but the analysis of the whole scene that led speakers to vary their proportion of active or passive sentences.

After this step, a structural scaffold is created; this is centered on the HN, the most dominant element. This step is clear in Japanese, a head-final language. After the initial structural scaffolding, lexical retrieval starts, with name-related gazes, that is, gazes directed to search for the corresponding lemmas in speech order. In Japanese, this process gives rise to a shift from the HN to the first mentioned element before SO, to shift again back to the HN (the second mentioned element) thereafter. Thus, in Japanese, an initial separation of structural scaffolding and lexical retrieval is observed.

On the other hand, in Spanish a more interwoven planning process seems to take place from the beginning, since the most dominant element (the HN) is also the first element to be mentioned. The preparation of this element takes until speech onset for Spanish speakers, with no shifting of gazes to different elements up to that point. In both cases, the structural scaffold is not completely set and fixed beforehand. Active and passive RCs show differences due to different processing costs, and those appear only after lexical retrieval has started.

Despite these initial differences found between Japanese and Spanish (with Japanese showing a structural bias, and Spanish allowing an interwoven planning of structural and lexical information), bilingual speakers do not show problems adapting to the requirements of the L2. This group of speakers showed signs of undergoing an initial apprehension stage, followed by an early structure scaffolding that did not differ from that showed by Japanese monolingual speakers. This resemblance, however, does not mean that they were following a conscious strategy of preparing the whole structure beforehand, and hence halting incremental production, as differences due to voice assignment suggested. The same differences as those found in monolinguals emerged, suggesting that bilingual planning unfolds incrementally, with speakers preparing a tentative scaffold beforehand. After that, lexical retrieval started giving way to the final representation of the utterance to be expressed. However, lexical retrieval was costly for bilingual speakers, showing long fixations to each of the elements in order. Somewhat surprisingly, there were almost no effects of the L1 grammar along the planning process, at least not in the first 1000 ms., while the structural scaffold was being constructed. Our results show converging evidence with Konopka & Forest's (2016) in that we observed that providing the verb beforehand helped participants to create the structural scaffold in exactly the same time window as monolingual speakers did, despite their problems with lexical retrieval. Additionally, bilingual speakers did not rely on animacy of single elements during the whole planning, but, just like monolinguals, on the conceptual prototypicality of the whole scene. They were able to flexibly adapt to the requirements of the L2 by focusing on the creation of a structural scaffold along an incremental process.

This study offers new and relevant evidence for the study of monolingual and bilingual language production. However, many are still the remaining questions it would be interesting to tackle in the near future. In the most immediate future, we are going to explore the time course of planning an RC with two important improvements: (1) responses will be elicited without using a question that focuses specifically on the element that corresponds the HN, which will ensure a more naturalistic planning; and (2) we will test RCs in two different positions in the main clause, namely, agent-modifying RC and patient-modifying RC's, which will allow us to explore the upper limits of the scope of RC planning and structural scaffold prior to lexical retrieval. Moreover, we are going to add more participants to our bilingual group to confirm the observed trends, and analyze in more detail the effects of lexical retrieval, voice assignment and the provision of a verb beforehand.

With these explorations, we hope more doors will open to the analysis of sentence planning, both from a monolingual and from a bilingual perspective. From a monolingual perspective, controlling the voice assigned to the RCs would help us explore how structural relations are finally established after lexical retrieval has started, without the confounding effects of animacy both in spoken responses and in gaze patterns. It also opens the possibility of exploring languages in which both HN-initial and HN-final RCs are allowed, as is the case of Tagalog. This would allow us to test whether the same flexibility encountered between languages is observed within the same linguistic system, by the very same speakers.

Finally, from a bilingual perspective, exploring the relation between L1 and L2 would allow us to assess models of bilingual sentence production, so as to attain a clearer picture of the time-course of planning in an L2. For that reason, we think that a structural priming experiment, paired with an eye-tracking methodology, will help us understand whether there are relations between L1 and L2 in structures with reversed word order across the two languages, and find out whether L1 interference occurs when both word orders match. Moreover, it is important to recall that we explored the time-course of planning of advanced bilinguals who reside in an L2-speaking country. Analyzing bilinguals with different levels of proficiency will provide further insights concerning the way the bilingual production system is established, and what are the connections between languages, as well as the idiosyncrasies that characterize learners' interlanguage.

In conclusion, this dissertation provides evidence for a structurally guided planning in RC production, even when the HN, the syntactically most dominant element of the utterance to be expressed, is located at the end of the clause. It provides converging evidence about the construction of a general scaffold previous to lexical retrieval in order and a wide scope of RC planning, encompassing the whole clause. However, it also incorporates evidence that grammatical function assignment is taking place and being established after lexical retrieval has started, supporting the view that early structural construction is not completely set at the point of lexical retrieval and allows for flexibility coming from unexpected sources during that process. Our results also shows how bilingual speakers are able to incrementally plan in their L2, flexibly adjusting their planning strategies to the requirements of its grammar, and showing, within the same participants and across different languages, how planning can efficiently move from relational to non-relational information in a single speaker in her two languages. It also opens the door to take a deeper look into the time-course of planning sentences from a different perspective, a perspective we are willing to explore still further.

Conclusiones

El desarrollo de medidas que permiten rastrear y registrar los movimientos oculares (*eye-tracking*) mientras se habla ha brindado la posibilidad de adquirir un conocimiento en mayor profundidad de la forma en que tiene lugar la planificación lingüística a lo largo del tiempo. Esta tesis se ha centrado en la exploración de algunas cuestiones aún no resueltas con respecto a la información que se emplea en la planificación de oraciones en hablantes monolingües y bilingües y en relación con el desarrollo temporal de los procesos que la integran. En los tres experimentos presentados en este trabajo, examinamos el proceso de planificación de oraciones de relativo (OR), primero desde una perspectiva translingüística, a través de la comparación entre hablantes monolingües de japonés y español, y posteriormente, en la producción de oraciones por parte de hablantes bilingües, con un grupo de bilingües tardíos español-japonés.

Nuestro objetivo principal era comparar dos modelos alternativos de la planificación en la producción de oraciones, los modelos de "incrementalidad jerárquica" e "incrementalidad lineal". Esta comparación permitía analizar cómo la información relacional (estructural) y no relacional (léxica) se usa a lo largo del proceso de planificación en una lengua de núcleo final y en una de núcleo inicial. La principal hipótesis era que un tipo de planificación del habla que da prioridad a las relaciones estructurales daría lugar a resultados similares en ambas lenguas, a pesar de las diferencias lineales (en el orden de palabras) entre ambas. En cambio, una planificación basada en los elementos léxicos mostraría diferencias translingüísticas desde el comienzo. Adicionalmente, teníamos como objetivo explorar cómo los hablantes bilingües de estas dos lenguas, que tienen orden de palabras opuestos, planifican el habla en la L2.

Merece la pena volver a traer y resumir en este punto varias de las conclusiones a las que se ha podido llegar a lo largo de los tres estudios. En primer lugar, cuando los hablantes se enfrentan a la tarea de producir una oración (ya sea para responder a una pregunta, como fue en el caso de nuestros estudios, o por motivaciones internas), crean inicialmente un mensaje no-lingüístico que contiene una idea general de lo que se pretende expresar. Esta codificación global (de la escena) mostró no estar influida por la "animacidad" de elementos léxicos aislados, pero sí, en cierta medida, por la distribución de la animacidad de la escena en su totalidad, ya que esta distribución representa la prototipicidad del evento representado en el dibujo. En nuestros resultados apenas hubo efectos de la animacidad en la proporción de fijaciones oculares a lo largo de todo el proceso. No obstante, las distintas combinaciones de animacidad modularon la cantidad de respuestas activas o pasivas en el caso concreto de las OR con el paciente como núcleo, lo que, una vez más, sugiere que es el análisis de la escena completa lo que lleva a los hablantes a variar sus respuestas entre oraciones activas y pasivas, y no tanto la accesibilidad de un elemento léxico concreto.

Tras este primer paso, se crea un esqueleto de la estructura (un andamio que sustenta la planificación): este esqueleto está centrado en el núcleo de la OR, que es el elemento más dominante (el que ocupa una posición más alta en la estructura sintáctica). Este proceso no ofrece dudas en el caso del japonés, una lengua de núcleo final. Después de construir el andamiaje inicial, comienza el acceso al léxico, lo que se observa por la aparición de miradas relacionadas con los nombres de objetos y personajes: se trata de miradas hipotéticamente dirigidas a buscar el lema correspondiente en el orden que ocupará en la cadena del habla. En japonés, este proceso da lugar a un desplazamiento de la mirada desde el núcleo hasta el elemento que será mencionado en primer lugar antes del comienzo del habla, para después volver a desplazarse al núcleo (elemento que será mencionado en segundo lugar en japonés). De este modo, en japonés el patrón de fijaciones oculares revela una separación inicial entre la creación de un esqueleto estructural y el acceso al léxico.

Por otro lado, en español la separación entre ambos procesos no está tan clara y parece observarse un proceso de planificación más interrelacionado desde el principio, dado que el elemento dominante (el núcleo de la OR) también es el primer elemento que se menciona. La preparación de este elemento se prolonga hasta el comienzo del habla en el caso de los hablantes de español, sin que hasta este punto haya ningún movimiento de la mirada entre elementos. En cualquier caso, en ninguna de las dos lenguas el esqueleto estructural queda completado y fijado de antemano. Las OR en voz activa y en voz pasiva muestran diferencias en los patrones oculares debidas a diferencias en la carga de procesamiento, aunque éstas aparecen sólo después de que el acceso al léxico haya comenzado.

A pesar de las diferencias iniciales encontradas entre el japonés y el español (recordemos que el japonés muestra un sesgo estructural y el español permite una planificación más entrelazada entre la información léxica y estructural), los hablantes bilingües no muestran problemas de adaptación a los requisitos de su L2 (i.e. el japonés). Este grupo de hablantes evidenció un proceso caracterizado por un entendimiento inicial de la escena, seguido de un andamiaje estructural temprano, el cual no difirió del mostrado por los hablantes japoneses. Además, los hablantes bilingües también mostraron diferencias debidas a la voz, las cuales tenían lugar después de que comenzara el acceso al léxico. Estos resultados, tomados en su conjunto, permiten descartar la posibilidad de que la semejanza inicial encontrada entre bilingües y monolingües japoneses (esto es, la planificación de la estructura centrada en el núcleo de la OR) sea debida a que los bilingües estén optando por una estrategia consciente de preparación de toda la estructura de antemano, lo que supondría renunciar a un proceso incremental en la producción de la L2. Más bien, parece que en el caso de los bilingües, la planificación también se desarrolla incrementalmente, con la creación inicial de un esqueleto tentativo de la estructura. Tras esta etapa, el acceso al léxico tiene lugar, dando paso a la representación final de la oración que será expresada. Sin embargo, hay indicios de que el acceso al léxico resultaba costoso para los hablantes bilingües, como muestran las largas fijaciones a cada uno de los elementos del dibujo en el orden en que iban a ser producidos. Para nuestra sorpresa, apenas hubo efectos de la gramática de la L1 a lo largo del proceso de planificación, al menos no durante los primeros 1000 ms., es decir, mientras el andamiaje de la estructura era creado. Nuestros resultados son consistentes con los datos obtenidos por Konopka y Forest (2016): así, en nuestro estudio observamos que el hecho de proveer a los participantes con el verbo de antemano les ayudó a crear el andamiaje de la estructura exactamente en el mismo intervalo temporal que a los monolingües japoneses, a pesar de sus problemas de acceso al léxico. Por otra parte, los hablantes bilingües no se vieron influidos por la animacidad de elementos individuales durante todo el desarrollo de la planificación. Sin embargo, y al igual que los hablantes monolingües, la distribución de la animacidad sí parece haber generado un efecto de la prototipicidad conceptual de toda la escena. En suma, los hablantes bilingües fueron capaces de adaptarse de manera flexible a los requisitos de la L2, dando prioridad a la creación de un andamiaje de la estructura dentro de un proceso incremental.

Este estudio ofrece datos nuevos y relevantes para la investigación de la producción del lenguaje por parte de hablantes monolingües y bilingües. Sin embargo, aún son muchas las preguntas que quedan sin responder y que creemos sería interesante abordar en un futuro cercano. De forma inmediata, nuestro siguiente paso consistirá en explorar el desarrollo temporal de la planificación de OR con dos mejoras importantes: (1) las respuestas serán elicitadas sin hacer uso de preguntas, ya que éstas sesgan la atención hacia uno de los elementos en concreto, en este caso el núcleo; de este modo, aseguraremos una planificación más natural y; (2) con este procedimiento se podrá poner a prueba ORs en dos posiciones diferentes de la cláusula principal: ORs que modifiquen al agente y ORs que modifiquen al paciente. Esta estrategia metodológica permitirá, a su vez, explorar los límites superiores del alcance de la planificación de las ORs y, con ello, del andamiaje estructural que es posible construir previamente al acceso al léxico. Por otro lado, un siguiente paso en nuestro estudio con bilingües es añadir más participantes, lo que nos permitirá confirmar las tendencias observadas y analizar con más detalle los efectos del acceso al léxico, la asignación de la voz y el suministro previo del verbo en la tarea de producción.

Con estas novedades, esperamos abrir otros caminos al análisis de la planificación de oraciones, desde la perspectiva de la producción tanto en hablantes monolingües como en hablantes bilingües. Desde el punto de vista de la producción en hablantes monolingües, un posible paso futuro sería controlar la voz asignada a las ORs, lo cual ayudaría a explorar de qué modo se establecen las relaciones estructurales una vez iniciado el acceso al léxico, controlando a la vez los efectos de la animacidad, tanto en las respuestas habladas como en los patrones de fijación ocular. Otra posibilidad es explorar lenguas que permitan la producción tanto de ORs de núcleo inicial como de núcleo final en la misma lengua, como es el caso del tagalo. Esto nos permitiría analizar si la misma flexibilidad que encontramos entre lenguas se observa dentro del mismo sistema lingüístico y por los mismos hablantes.

Finalmente, desde la perspectiva de la producción en bilingües, consideramos que un paso importante consistiría en explorar las relaciones entre la L1 y la L2. Esto nos permitiría poner a prueba los distintos modelos de producción de oraciones en hablantes bilingües y tener una imagen más clara del desarrollo temporal de la planificación en la L2. Por este motivo, consideramos que un experimento de *priming* estructural, acompañado de un registro de movimientos oculares, ayudaría a entender si la L1 y la L2 están conectadas en estructuras en las que el orden es el opuesto en las dos lenguas, y observar si existe interferencia de la L1 cuando el orden de palabras es el mismo. Por otra parte, es importante recordar que en nuestro estudio analizamos el desarrollo temporal de la planificación de ORs en bilingües de nivel avanzado, que además residían en el país donde la L2 era lengua nativa. Así, un análisis de la producción de hablantes bilingües con distintos niveles de competencia en la L2 arrojaría evidencia muy valiosa sobre la forma en la que el sistema de producción de oraciones se consolida en estos hablantes, así como sobre las conexiones entre lenguas y la idiosincrasia que caracteriza la interlengua de los aprendices.

En conclusión, este trabajo aporta evidencia de una planificación de las ORs que está guiada estructuralmente, incluso cuando el núcleo, el elemento sintácticamente dominante, está situado al final de la cláusula. Aporta evidencia convergente acerca de la creación de un andamiaje estructural general, previo al acceso léxico (el cual se da en el mismo orden de producción), así como evidencia de un alcance amplio en la planificación de las ORs, que abarca toda la cláusula. Sin embargo, nuestro trabajo también proporciona indicios de que la asignación de la función gramatical no queda enteramente concluida hasta el punto en el que comienza el acceso al léxico, lo que facilita la suficiente flexibilidad para dar cabida a elementos inesperados durante la planificación. Adicionalmente, nuestros resultados muestran cómo los hablantes bilingües son capaces de planificar incrementalmente en su L2, mediante un ajuste de sus estrategias de planificación a los requisitos de la gramática de la L2. Estos resultados muestran cómo la planificación puede moverse eficientemente desde la información relacional a la no-relacional en los mismos hablantes en sus dos lenguas. A través de estos resultados se abre la puerta a un análisis en profundidad del desarrollo temporal de la planificación de oraciones bajo una perspectiva diferente, perspectiva que seguiremos explorando en el futuro.

References

Allum, P.H. & Wheeldon, L.R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol 33(4),* 791-810.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36*, 233–240.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173-218). New York: Cambridge University Press.

Beauvillain, C. & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language, 26*, 658-672.

Bernolet, S. (2008). *Lexical-syntactic representations in bilingual sentence production.* PhD dissertation. University of Gent, Belgium.

Bernolet, Sarah, Hartsuiker, R., & Collina, S. (2016). The persistence of syntactic priming revisited. *Journal of Memory and Language*, *91*, 99–116.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33, 931-949.*

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2013). From language-specific to shared syntactic representa- tions: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition, 127*(3), 287–306.

Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clause in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology, 62,* 1915–1929.

Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Essex: Pearson Education.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355-387.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. Journal of Memory and Language, 31, 99–127.

Bock, J.K. ; Dell, G. ; Chang, F. & Onishi, K. (2007). Persistent structural priming from language comprehension to language production. Cognition, 104 (3):437-458.

Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129*, 177-192.

Bock, J. K., & Levelt, W. (1994). Language production: Grammatical encoding. In Gernsbacher, Morton Ann (Ed). (1994), *Handbook of Psycholinguistics*, pp. 945-984. San Diego, CA, US: Academic Press, Inc

Bock, J. K. & Ferreira, V. S. (2014). Syntactically speaking. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds), *The Oxford Handbook of Language Production* (pp. 21-46). Oxford: Oxford University Press.

Bock, J. K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. M. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world,* 249-278. New York: Psychology Press.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language, 48(4),* 653-685.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition* 35, 1–39.

Bock, J. K., Loebell, H. & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150-171.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45-93

Bock, J. K. & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47—67.

Boersma, P., Weenink, D. 2014. *Praat: doing phonetics by computer* [Computer program]. Version 5.3.71, retrieved 9 April 2014 from http://www.praat.org/

Bosque, I. & Demonte, V. (1999). *Gramática descriptiva de la lengua española*. [Descriptive grammar of the Spanish language] Madrid: Colección Nebrija y

Bello, Espasa.

Branigan, H. P. & Feleki, E. (1999). Conceptual accessibility and serial order in Greek language production. In M. Hahn & S. C. Stoness (Eds), *Proceedings of the 21st Conference of the Cognitive Science Society* (pp. 96-101). Mahwah: Erlbaum.

Branigan, H. P., Pickering, M. J. & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonometric Bulletin & Review, 6(4)*, 635-640.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13–B25.

Branigan, H.P., Pickering, M.J., Stewart, A.J. & Mclean, J.F. (2000). Memory & Cognition, 28 (8): 1297-1302.

Branigan, H.P., Pickering, M.J, & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. Lingua, 118, 172-189.

Broersma, M. (2005). *Phonetic and Lexical Processing in a Second Language*. Nijmegen, The Netherlands: PhD dissertation, Radboud University Nijmegen.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition, 109(2),* 274-280.

Brown-Schmidt, S. & Konopka, A.E. (2015). Processes of incremental message planning during conversation. *Psychonomic Bulletin and Review*, *22*(3), 833-843.

Brown-Schmidt, S. & Tanenhaus, M.K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. Journal of Memory and Language, 54, 592-609.

Cai, Z.G., Pickering, M.J. & Branigan, H.P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language, 66*, 833–849.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Brain and Behavioral Sciences, 22*, 77–126.

Carreiras, M., & Clifton, C., Jr. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, 36, 353-372.

Carreiras, M., & Clifton, C., Jr. (1999) Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & Cognition*, 27, 826-833.

Carreiras, M., Duñabeitia, J.A., Vergara M., de la Cruz-Pavía, I & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition, 115,* 79-92.

Carreiras, M., Salillas, E., & Barber, H. (2004). Event related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research, 20, 98-105.*

Chang, F., Dell, G.S., Bock, J.K. & Griffin, Z.M. (2000) Structural priming as implicit learning: A comparison of models of sentence production. Journal of Psycholinguistic Research, 29(2): 217–229.

Chang, F. (2002) Symbolically speaking: A connectionist model of sentence production. *Cognitive Science, 26*: 609–651.

Chang, F., Dell, G.S. & Bock, J.K. (2006) Becoming syntactic. Psychological Review, 113: 234–272.

Chen, B., Ning, A., Bi, H. & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than object-relative clauses. *Acta Psychologica, 129*(1): 61–65

Christiansen, M.H & Chater, N. (2016). *Creating language: integrating evolution, acquisition, and processing*. Cambridge, MA: The MIT Press.

Clahsen H. & Felser, C. (2006a). How native-like is non-native language processing? *Trends in Cognitive Science, 10,* 564–570.

Clahsen H. & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition, 24*(1), 94-102.

Comrie, B. (2002). Typology and language acquisition: The case of relative clauses. In A. G. Ramat (Ed.), *Typology and second language acquisition* (pp. 19-37). Berlin: Mouton de Gruyter.

Comrie, B. (2007). The acquisition of relative clauses in relation to language typology. *Studies in Second Language Acquisition*, *29*, 301–309.

Comrie, B., & Keenan, E. (1979). Noun Phrase Accessibility Revisited. *Language, 55*(3), 649-664.

Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365–397.

Costa, A., & Santesteban, M. (2004a). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491–511.

Costa, A., & Santesteban, M. (2004b). Bilingual word perception and production: two sides of the same coin?. *Trends in Cognitive Sciences. 8,* 253.

Cuetos, F. & Mitchell, D.C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, *30,* 73-105.

De Bot, K. (1992) A Bilingual Processing Model: Levelt's 'Speaking' Model Adapted. *Applied Linguistics, 13*, 1-24.

De Groot, A. M. B. (2011). *Language and Cognition in bilinguals and multilinguals. An introduction.* New York: Psychology press.

De Jong N. H., Steinel M. P., Florijn A. F., Schoonen R., Hulstijn J. H. (2012). The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In Housen A., Kuiken F., Vedder I. (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 121–142). Amsterdam: John Benjamins

Dekydtspotter, L., Schwartz, B.D., & Sprouse, R.A. (2006). The Comparative Fallacy in L2 Processing Research. In M.G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 33-40). Somerville, MA: Cascadilla Proceedings Project.

Del Río, D. (2008). *Asignación de papel temático y comprensión de oraciones con extracción de sujeto y objeto en español. Un estudio sobre las demandas de procesamiento a nivel cognitivo y neurofisiológico*. [Assignment of thematic role and comprehension of sentences with subject and object extraction in Spanish. A study about processing demands to a cognitive and neurophysiological level]. Unpublished doctoral thesis. Faculty of Psychology. Universidad Complutense de Madrid, Spain.

Dell, G.S. & Ferreira, V.S. (2016). Thirty years of structural priming: An introduction to the special issue. *Journal of Memory and Language, 91*, 1-4.

Demonte, V. & Fernández Soriano, O. (2007). "La periferia izquierda oracional y los complementantes del español". [The left sentente periphery and Spanish complementizers]. In J. Cuartero Otal & M. Emsel (Eds.). *Vernetzungen: Bedeutung in Wort, Satz und Text.* Pieterlen: Peter Lang.

Desmet, T., & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language, 54*, 610–632.

Dijkstra, T. (2007) The multilingual lexicon. In M.G. Gaskell (Ed.). *The Oxford Handbook of Psycholinguistics* (pp. 251-265). Oxford: Oxford University Press.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and language, 62*(3), 284-301.

Dijkstra, T., & van Heuven, W.J.B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175-197.

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish-English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica, 128*, 501-513.

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research*, 26, 443-472.

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition, 10*, 101-116. Felser & Roberts, 2007

Felser, C., Roberts, L., Marinis, T. & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics 24*, 453–489.

Fernández Lagunilla, M. & Anula, A. (1995). *Sintaxis y cognición. Introducción al conocimiento, el procesamiento y los déficits sintácticos* [Syntax and cognition: An introduction to syntactic knowledge, processing and deficits]**.** Madrid: Síntesis.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30,* 210 –233.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84.

Ferreira, V. S. (2010). Language production. *Wiley Interdisciplinary Reviews: Cognitive Science, 1(6)*, 834–844.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*(7-8), 1011–1029.

Ferreira, V.S., Bock, K., Wilson, M.P. & Cohen, N.J. (2008). Memory for syntax despite amnesia. *Psychological Science, 19*(9), 940-946.

Ferreira, V.S. & Slevc, L. R. (2007). Grammatical encoding. In M.G. Gaskell (Ed.). *The Oxford Handbook of Psycholinguistics* (pp. 453-469). Oxford: Oxford University Press.

Foucart, A. & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory & Language*, 66226–248.

Frazier, L. 1979. *On Comprehending Sentences: Syntactic parsing strategies.* Ph.D. Dissertation, University of Connecticut.

Frazier, L., & Clifton Jr, C. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research, 26*(3), 277-297.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In Heredia, R. R. & Altarriba, J. (Eds.) *Bilingual sentence processing* (pp. 217-236). North Holland: Elsevier.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21*, 175–198.

Fromkin, V. (1973). Speech errors as linguistic evidence. The Hague: Mouton.

Ganushchak, L. Y., Konopka, A. E., & Chen, Y. (2014). What the eyes say about planning of focused referents during sentence formulation: a cross-linguistic investigation. Frontiers in Psychology, 5, 1124.

Garrett, M.F. (1975). The analysis of sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 9). London: Academic Press.

Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed). *Language Production* (Vol. 1). London: Academic Press.

Garrett, M.F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions.* London: Academic Press.

Gass, S. M. & Selinker, L. (2008). *Second Language Acquisition. An Introductory Course (3ª ed.)*. New Jersey: Lawrence Erlbaum Associates.

Gennari, S. P. & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58,* 161-187.

Gennari, S. P. & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition, 111*, 1-23.

Gennari, S. P., Mirkovic, J. & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology, 65,* 141-176.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition 68*, 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, W. O'Neil (Eds.), *Image, language, brain.* Cambridge, MA: MIT Press, 95–126.

Gibson, E. & Wu, H.I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes, 28*, 125–155.

Gleitman, L. R., January, D., Nappa, R. & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language 57,* 544–569.

Gollan, T.H., Bonanni, M.P., & Montoya, R.I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip-of-the-tongue equally often. *Neuropsychology, 19,* 278–287.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814.

Gordon P.C., Hendrick, R. & Johnson, M. (2001). Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory and Cognition, 27(6), 1411-1423.

Green D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82,* B1-B14.

Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review, 10,* 603–609.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. Henderson & F. Ferreira, (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 213-247). New York: Taylor and Francis.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11,* 274-279.

Gutiérrez-Bravo, R. (2003). Subject inversion in Spanish Relative Clauses: A case of prosody-induced word order variation without narrow focus. In T. Geerts & H. Jacobs (Eds.). *Romance Languages and Linguistic Theory*, Amsterdam: John Benjamins.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4*, 123-141.

Hartsuiker, R. J., Beerts, S., Loncke, M., Desmet, T., & Bernolet, S. (2016). Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of Memory and Language*, *90*, 14–30.

Hartsuiker, R.J., Bernolet, S., Schoonbaert, S., Speybroeck, S. & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. Journal of Memory & Language, 58, 214–238.

Hartsuiker, R. J., & Kolk, H.H.J. (1998). Syntactic facilitation in agrammatic sentence production. In: Brain and Language 62, 221-254.

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica, 128,* 479-489.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science*, 15, 409-414.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75*, B27-B39.

Haskell, T.R. & MacDonald, M.C. (2005). Constituent structure and linear order in language production: Evidence from subject-verb agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 891–904.

Hatzidaki, A., Branigan, H.P., & Pickering, M.J. (2011). Co-activation of syntax in bilingual language production. *Cognitive Psychology* 62, 123–150.

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. Crocker, M. Pickering and C. Clifton, jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 259-281). Cambridge: Cambridge University Press

Hirose, Y. (2006). Processing relative clauses in Japanese: coping with multiple ambiguities. In M. Nakayama, R. Mazuka, Y. Shirai & P. Li (Eds.), *The handbook of East Asian Psycholinguistics* (Vol. II: Japanese) (pp. 264-269). Cambridge: Cambridge University Press.

Hoshino, N., Dussias, P.E., & Kroll, J. F. (2010). Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, *13*, 87-98.

Hoshino, N. & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106,* 501–511.

Hsiao, F. & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90,* 3-27.

Huang, J., Pickering, M.J., Yang, J., Wang, S. & Branigan, H.P. (2016). The independence of syntactic processing in Mandarin: Evidence from structural priming, *Journal of Memory and Language, 91,* 81-98.

Huddleston, R. & Pullum, G.K. (2002). *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.

Hwang, H. & Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40, 1363-1376.

Igoa, J.M, & García-Albea, J.E. (1999). Unidades de planificación y niveles de procesamiento en la producción del lenguaje [Planning units and level of processing in language production]. In M. de Vega & F. Cuetos (Eds.), Psicolingüística del español (pp. 375-468). Madrid: Editorial Trotta.

Iori, I. (2001). 新しい日本語学入門。言葉の仕組みを考える。[*Atarashii Nihongo-gaku Nyuumon. Kotoba-no Shikumi-o kangaeru.* New introduction to the study of Japanese. Thinking in the constrcution of words]. Tokio: Surie Network.

Ishizuka, T. (2005). Processing relative clauses in Japanese. *UCLA Working Papers in Linguistics 13*(2). 135−157.

Ishizuka, T., Nakatani, K., & Gibson, E. (2003, March). Relative clause extraction complexity in Japanese. Poster presented at the 16th annual CUNY conference on human sentence processing, Massachusetts Institute of Technology, Cambridge, MA.

Ishizuka, T., Nakatani, K., & Gibson, E. (2006, March). Processing Japanese relative clauses in context. Paper presented at the 19th annual CUNY conference on human sentence processing, CUNY, New York.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288.

Ivanova, I., Pickering, M.J., Branigan, H.P., McLean, J.F. & Costa, A. (2012). The comprehension of anomalous sentences: evidence from structural priming. *Cognition, 122*(2), 193-209.

Iwasaki, S. (2002). *Japanese*. Amsterdam: John Benjamins.

Jackson, C., & Roberts, L. (2010). Animacy affects the processing of subject–object ambiguities in the second language: Evidence from self-paced reading with German second language learners of Dutch. *Applied Psycholinguistics, 31*(4), 671-691.

Jäger, L., Benz, L., Roeser, J., Dillon, B. & Vasishth, S. (2015). Teasing apart retrieval and encoding interference in the processing of anaphors. *Frontiers in Psychology,* 6(506).

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21*(2), 121–151.

Juffs, A., & Rodríguez, G.A. (2015). *Second language sentence processing*. New York: Routledge.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Kahraman, B., Sato, A., Ono, H. & Sakai, H. (2010). Relative clauses processing before the head-noun: Evidence for strong forward prediction in Turkish. In H. Maezawa & A. Yokogoshi (Eds.), *MIT Working Papers in Linguistics* 61, 155−170. Cambridge, MA: MITWPL.

Kahraman, B. (2012). Sentence processing of nominative-genitive conversion in Japanese by Turkish speaking learners and native speakers. In Y. Otsu (Ed.), *The Proceedings of Thirteenth Tokyo Conference on Psycholinguistics.* 81−102. Tokyo: Hituzi Syobo.

Kahraman, B. & Sakai, H. (2015). Relative clause processing in Japanese: Psycholinguistics Investigation into typological differences. In M. Nakayama (Ed.), *Handbooks of Japanese Language and Linguistics* (pp. 423-456) Berlin: De Gruyter Mouton.

Kanno, K. (2007). Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition, 29, 197-218.*

Kantola, L., & van Gompel, R. P. G. (2011). Between- and within-language priming is the same: Evidence for shared bilingual syntactic representations. *Memory & Cognition*, *39*(2), 276–290.

Kaschak, M.P. (2007) Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition, 35*, 925–937.

Kaschak, M.P., Kutta, T. J. & Coyle, J. (2014). Long and short term cumulative structural priming effects. *Language, Cognition, and Neuroscience*, 29, 728–743.

Kaschak, M.P., Kutta, T.J. & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition, 39,* 381–388.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. Linguistic Inquiry, 8, 63-99.

Keenan, E. L & Comrie, B. (1979). Data on the Noun Phrase Accessibility Hierarchy. *Language, 55* (2), 333-351

Kellly, M. H., Bock, J. K., Keil, F. C. (1986). Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language, 25*, 59-74.

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*(2), 201-258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*(2), 185-209.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580-602.

King, J. W. and Kutas, M. (1995) Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376-395.

Klaus, J., Mädebach, A., Oppermann, F., & Jescheniak, J. D. (2017). Planning sentences while doing other things at the same time: effects of concurrent verbal and visuospatial working memory load. *The Quarterly Journal of Experimental Psychology*, *70*(4), 811–831.

Kobayashi, N., Niwa, J. & Yamamoto, H. (1996).日本語能力の新しい測定法 <SPOT>. [*Nihongo nouryoku-no atarashii sokutei houhou: SPOT* – A new method of measuring Japanese proficiency: SPOT]. *Sekai-no Nihongo kyouiku, 6*, 201-218.

Kolers, P. (1966). Reading and talking bilingually. *American Journal of Psychology*, *3*, 357–376.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language, 66,* 143–162.

Konopka, A. E. & Brown-Schmidt, S. (2014). Message encoding. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford Handbook of Language Production* (pp. 3-20). Oxford: Oxford University Press.

Konopka, A.E. & Kuchinsky, S. (2015). How message similarity shapes the timecourse of sentence formulation. *Journal of Memory and Language, 84*, 1-23.

Konopka, A. E. & Forest, T. A. (2016). Linguistic experience (L1 vs. L2) shapes sentence formulation. Talk at *The 29th Annual CUNY Conference on Human Sentence Processing*. University of Florida, Gainesville, Florida. 3-5 March.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology, 73*, 1-40.

Kroll, J.F., & Gollan, T.H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds), *The Oxford Handbook of Language Production* (pp. 165-181). Oxford: Oxford University Press.

Kubo, T. (2016). OS 語順の産出における普遍性と言語個別性—カクチケル語VOS 語順における検討— [*OS-gojun-no sanshutsu-ni okeru fuhensei to gengo kobetsusei: kakuchikeru-go VOS-gojun-ni okeru kento*, Linguistic universal and idiosyncrasies in the production of OS word orders: An analysis of VOS word order in Kaqchikel]. Unpublished PhD. dissertation. Hiroshima University, Hiroshima, Japan.

Kuchinsky, S. E., Bock, K. & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37* (3), 748-756.

Kwon, N., Kluender, R., Kutas, M. & Polinsky, M. (2013). Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. Language, 89(3), 537-585.

Kwon N., Polinsky M. & Kluender R. (2006). Subject preference in Korean. In D. Baumer, D. Montero & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics (WCCFL 25)* (pp. 1–14). Somerville, MA: Cascadilla Proceedings Project.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J.F. Kroll & A.M.B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289–307). Oxford: Oxford University Press.

Lee, E. K., Brown-Schmidt, S., & Watson, D. G. (2013). Ways of looking ahead: Hierarchical planning in language production. *Cognition*, *129*(3), 544–562.

Levelt, W. J. M (1989) *Speaking: from intention to articulation.* Cambridge: MIT Press.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

Lewis, R.L. & Vasishth, S. (2005). An Activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29,* 375–419.

Lin, Y. & Garnsey, S. (2011). Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In L. Frazier, T. Roeper, and K. Wexler (Eds.), *Studies in Theoretical Psycholinguistics: Vol. 38. Processing and Producing Head-Final Structures* (pp. 241–275). NY: Springer.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883.

Litcofsky, K.A., Tanner, D., & van Hell, J.G. (2016). Effect of language proficiency, use, and cognitive function on bilingual word production and comprehension. *International Journal of Bilingualism, 20 (6),* 666-683.

Loebell, H. & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*(5), 791 - 824.

MacDonald, M. C. (2013a). Inviting Production to the cognitive basis party. In M. Sanz, I. Laka & M. K. Tanenhaus (Eds.), *Language Down the Garden Path: The Cognitive and Biological Basis of Linguistic Structures.* Oxford: Oxford University Press.

MacDonald, M. C. (2013b). How language production shapes language form and comprehension. *Frontiers in Psychology, 4* (226).

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676-703.

Macnamara, J. (1967). The Bilingual's Linguistic Performance-A Psychological Overview. *Journal of Social Issues*, *23*(2), 58–77.

MacWhinney, B. (1997). Second language acquisition and the Competition Model. In J. Kroll & A. De Groot (Eds.), *Tutorials in bilingualism*. Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2001). Last words. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 257-264). Amsterdam: John Benjamins.

MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition, 29,* 95-141.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language, 91,* 5-27.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*, 50–68.

Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language, 54*, 466–490

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55,* 381-401.

McDonald, J. L., Bock, K. & Kelly, M. H. (1993). Word and world order: Semantic, Phonological and metrical determinants of serial order. *Cognitive Psychology, 25,* 188-230.

Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition*, *23*(4), 477–494.

Meijer, P. J. A., & Fox Tree, J. E. (2003). Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology, 50*, 184–195.

Messenger, K., Branigan, H.P., McLean, J.F. & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language, 66,* 568-587.

Meuter, R.F.I. & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language 40*, 25–40 .

Meyer, A.S. (1996). Lexical access in phrase and sentence production: results from picture-word interference experiments. *Journal of Memory and Language, 35*, 477–496.

Meyer, A.S. (2004). The Use of Eye Tracking in Studies of Sentence Generation. In J.M. Henderson and F. Ferreira (Eds), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 191-211). New York, NY, US: Psychology Press.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25–B33.

Miyamoto, E.T., & Nakamura, M. (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. In G. Garding & M. Tsujimura (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*. (pp. 342–355). San Diego, CA.

Momma, S., Slevc, L. R., & Phillips, C. (2015). The Timing of Verb Selection in Japanese Sentence Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 813-882.

Montag, J. L. & MacDonald, M. C. (2009). Word order doesn't matter: Relative clause production in English and Japanese. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 2594-2599). Cognitive Science Society. Amsterdam, Holland.

Myachykov, A., & Tomlin, R. (2008). Attention-capturing priming and structural choice in Russian sentence production. *Journal of Cognitive Science, 6*(1), 31–48.

Norcliffe, E., & Konopka, A. E. (2015). Vision and language in cross-linguistic research on sentence production. In R. K. Mishra, N. Srinivasan, & F. Huettig (Eds.), *Attention and vision in language processing* (pp. 77-96). New York: Springer.

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, *30*(9), 1187–1208.

O'Grady, W. (1997). *Syntactic Development*. Chicago, IL: University of Chicago Press.

Onishi K. H., Murphy G. L, Bock J. K. (2008). Prototypicality in sentence production.

Cognitive Psychology, 56, 103–141.

Ozeki, H. & Shirai, Y. (2007). Does the Noun Phrase Accesibility Hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition, 29, 169-196.*

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition 24*, 501–528.

Pickering MJ, Branigan HP. The representation of verbs: Evidence from syntactic priming in language production. Journal of Memory and Language. 1998, 39, 633–651.

Pickering, M. J., & Ferreira, V. S. (2008). Structural Priming: A Critical Review. *Psychological Bulletin, 134*(3), 427–459.

Poulisse, N., & Bongaerts, T. (1994). First language use in Second Language production. *Applied Linguistics, 15*(1), 36-57.

Prat-Sala, M. & Branigan, H.P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English & Spanish. *Journal of Memory & Language, 42*, 168-182.

Qiao X., Shen L. & Forster K. (2012). Relative Clause Processing in Mandarin: Evidence from the Maze Task. *Language and Cognitive Processes*, 27, 611–630.

Real Academia Española, Asociación de Academias de la Lengua Española (2009). *Nueva Gramática de la lengua Española* [New Grammar of Spanish]. *Vol. II.* Madrid: Espasa libros.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*, 1–23.

Roberts, L. (2013). Sentence processing in bilinguals. In R.P.G. van Gompel (Ed.), *Sentence processing*. (pp. 221-246). New York: Psychology press.

Rodrigo, L. (2013). Semantic cues in relative clause production by Spanish speakers learning Japanese. Unpublished Master thesis. Kobe City University of Foreign Studies, Kobe, Japan.

Rodrigo, L., Kubo, T., Tanaka, M, & Koizumi, M. (2014). The effects of word order accessibility in sentence production: the case of Kaqchikel language. Unpublished work.

Rodrigo, L., Tamura, A., Kubo, T., Tanaka, M, & Koizumi, M. (2016). Priming word order in bilingual speakers with free word order variation: the case of Spanish-Kaqchikel bilinguals. Unpublished work.

Rosen, S.T. (1999). The syntactic representation of linguistic events. *Glot international, 4*(2), 3-11.

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129*(2), 256–263.

Runnqvist, E., Fitzpatrick, I., Strijkers, K. & Costa, A. (2012). An appraisal of the bilingual language production system: Quantitatively or qualitatively different from monolinguals? In T.K. Bhatia & W.C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism, Second Edition.* (pp. 244-265). Chichester, UK: John Wiley & Sons, Ltd.

Salamoura, A. & Williams, J.N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual mental lexicon. *Applied Psycholinguistics, 28*, 627-660.

Sanz, M. (2005). Restructuring processes of Interlanguages: The Acquisition of Spanish COMP Structures by Japanese speakers. *Kobe-shi Gaikoku-go Daigaku, Gaikoku-gaku kenkyuu, 65, 1-20.*

Sanz, M. & Igoa, J.M. (2012). Syntax and sentence processing in relation to language teaching. In M. Sanz and J.M. Igoa (Eds.), *Applying Language Science to Language Pedagogy. Contributions of Linguistics and Psycholinguistics to Second Language Teaching*. Cambridge: Scholars Publishing.

Sasaki, Y. (1994). Paths of processing strategy transfers in learning Japanese and English as foreign languages. *Studies in Second Language Acquisition, 16,* 43-72.

Sato, A., Kahraman, B. & Sakai, H. (2010). Does frequency of occurrence make relative clause processing easier in Japanese? Poster Presented at *the 23rd Annual CUNY Conference on Human Sentence Processing*, New York University, New York, 16−20 March.

Sato, A., Kahraman, B. & Sakai, H. (2012). When is the object relative clause easier to process than the subject relative clause? *Technical Report of IECIE 112*(145). 41−46.

Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin Jr., R. D., & Levinson, S. C. (2013). Dependencies first: Eye tracking evidence from sentence production in Tagalog. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1265-1270). Austin, TX: Cognitive Science Society.

Schoonbaert, S., Hartsuiker, R. J. & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language 56*(2), 153-171.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology, 11*, 17–50.

Schriefers, H., Friederici, A., & Kühn, K. (1995). The processing of locally ambiguous relative clauses in German . *Journal of Memory and Language, 34*, 499-520.

Schriefers, H., Teruel, E., & Meinshausen, R. M. (1998). Producing Simple Sentences: Results from Picture–Word Interference Experiments. *Journal of Memory and Language*, *39*(4), 609–632.

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty and M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 382-408). Oxford: Blackwell Publishing.

Sekerina, I. (2003). The late closure principle in processing of ambiguous Russian sentences. In P. Kosta and J. Frasek (Eds.), *Current approaches to formal*

Slavic linguistics: Contributions of the second European conference on formal description of Slavic languages (FDSL II) held at Postdam University, November 20-22, 1997. (pp. 205-217). Frankfurt/M., Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics. 10*, 209-231.

Shin J.A. (2010). Structural priming and L2 proficiency effects on bilingual syntactic processing in production. *Korean Journal of English Language and Linguistics, 10*, 499–518.

Shin, J. A. & Christianson, K. (2009). Syntactic processing in Korean–English bilingual production: Evidence from cross-linguistic structural priming. Cognition 112 (2009) 175–180.

Shirai, Y. & Ozeki, H. (2007). Introduction. *Studies in Second Language Acquisition, 29, 155-167.*

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73,* 205–246.

Slevc, L.R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1503-1514.

Suzuki, Y. (2014). 第二言語の文法知識の自動化の簡易的な測定方法：WEB版SPOTとACTFL口頭能力測定(OPI)の比較。[*Dainigengo-no bunpou chishiki-no jidouka-no kaniteki-na sokutei houhou* – A practical method of measuring automaticity in Second Language grammar] *IEICE technical report, 114*(100), 49-54.

Swets, B., Desmet, T., Hambrick, D.Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General, 136*(1), 64-81.

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2008). Individual differences in the planning scope of language production. Paper presented at *the 49st meeting of the Psychonomic Society*, Chicago, IL (November).

Szmalec, A., Brysbaert, M., & Duyck, W. (2012). Working memory and (second) language processing. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism : theoretical and applied approaches* (pp. 74–94). Cambridge: Cambridge University Press.

Talama, A., Kroll, J.F., & Dufour, R. (1999). Form related errors in second language learning: A preliminary stage in the acquisition of L2 vocabulary. *Bilingualism: Language and Cognition, 2,* 45–58.

Tanaka, M., Branigan, H.P., McLean, J.F., & Pickering, M.J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language, 65,* 318-330.

Tanaka, M., Pickering, M. J. & Branigan, H. P. (2009). *Structural priming of word order and voice in Japanese and English.* Unpublished manuscript.

Traxler, M. J., Morris, R.K. & Seely, R.E. (2002). Processing Subject and Object Relative Clauses: Evidence from Eye Movements. *Journal of Memory and Language, 47*, 69-90.

Traxler, M. J, Williams, R. S., Blozis, S. A. y Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language, 53, 204-224.*

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162-189). Cambridge: Cambridge University Press.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules.* Cambridge MA: M.I.T. Press.

Tsujimura, N. (2007). *An Introduction to Japanese Linguistics* (2ª ed.). Oxford: Blackwell Publishing.

Ueno, M. and Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes, 23:* 5, *646-688.*

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative / procedural model. *Nature Reviews Neuroscience, 2*, 717–726.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative / procedural model. *Cognition. 92*, 231–270.

van de Velde, M., & Meyer, A. S. (2014). Syntactic flexibility and planning scope: the effect of verb bias on advance planning during sentence recall. *Frontiers in Psychology*, *5*, 1174.

van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing "easy" and "hard" events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*(1), 124–144.

Vasishth, S. (2015). *A meta-analysis of relative clause processing in Mandarin Chinese using Bias Modeling.* MS Dissertation submitted at the School of Mathematics and Statistics of the University of Sheffield.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese Relative Clauses: Evidence for the Subject-Relative Advantage. *PLOS ONE 8*(10): e77006

Veenstra, A., Acheson, D. J., Bock, K., & Meyer, A. S. (2014). Effects of semantic integration on subject–verb agreement: evidence from Dutch. *Language, Cognition and Neuroscience*, *29*(3), 355–380.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Cross-linguistic variability in subject–verb agreement errors. *Cognition, 61*, 261-298.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 423–440.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. Cognition, 85, 79–112.

Wheeldon, L., Ohlson, N., Ashby, A., & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *Quarterly Journal of Experimental Psychology, 66*(8), 1653–1673.

Williams, J.N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English *wh*- questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22,* 509-540.

Yamada, Y. (1995) (Ed.). 中級スペン文法 [*Chuu-kyuu Supein bunpoo,* Intermediate level of Spanish Grammar]. Tokyo: Hakusuisha.

Appendices

Appendix 1: List of experimental items

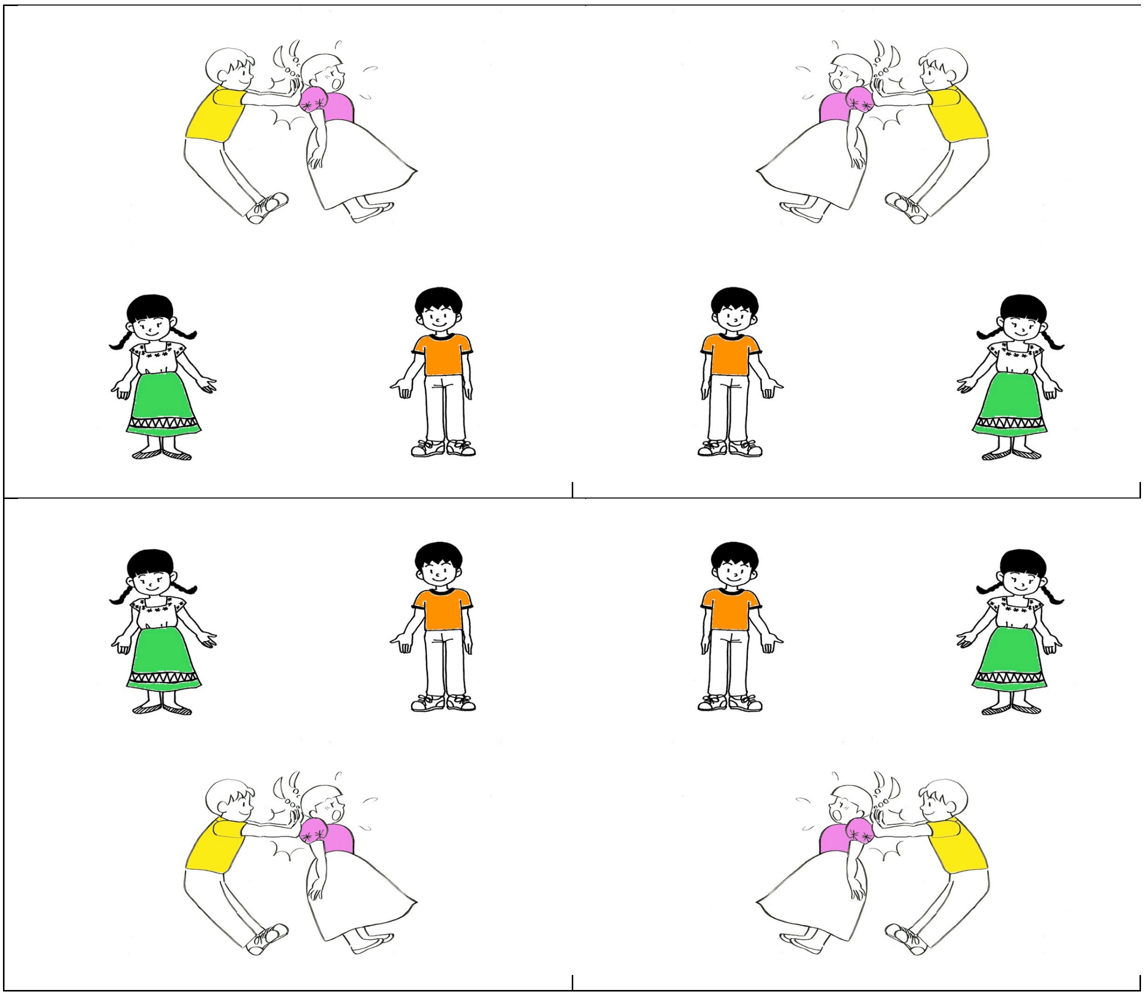
Experimental items include a total of 30, hand-drawn pictures presented with questions that were directed to either the agent or the patient of the event, so as to elicit:

1. RCs with the agent as HN
2. RCs with the patient as HN

There are ten different verbs, each used with three animacy combinations:

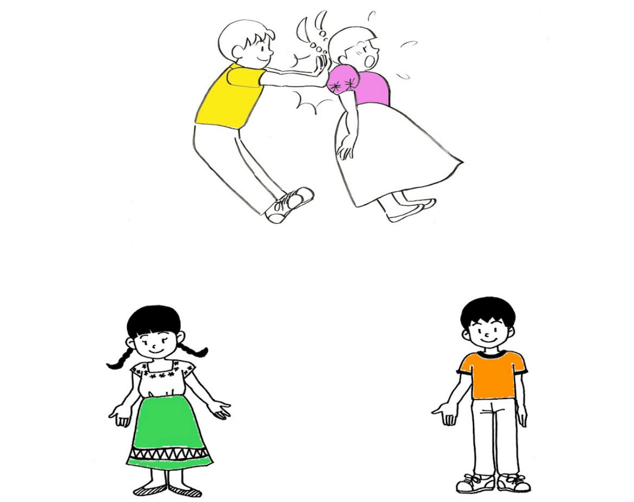
1. Animate agent – Animate patient
2. Animate agent – Inanimate patient
3. Inanimate agent – Animate patient

All pictures are presented in this order in this Appendix, each headed by the provided verb. All items were counterbalanced, presenting the agent and the patient in any of the four possible locations, as can be seen below. However, only the version with the agent in the upper left part is shown in the list of items presented in this Appendix.



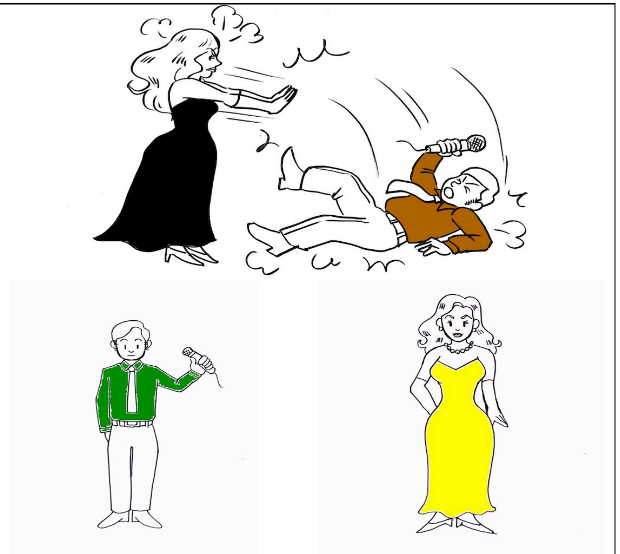
List of experimental Items with provided questions in Japanese and Spanish, and provided verb.

Verb 1: Empujar / 押す (Osu) – “Push”

1. Agent-HN: *a. Spanish:* ¿Quién lleva una camiseta amarilla? Who wears a T-shirt yellow *b. Japanese:* 誰が黄色いＴシャツを着ていますか？ (Dare-ga kiiroi T-shatsu-o kiteimasuka?) Who-NOM yellow T-shirt-ACC is wearing-Q → “Who is wearing a yellow T-shirt?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva una blusa rosa? Who wears a blouse pink *b. Japanese:* 誰がピンクのブラウスを着ていますか？ (Dare-ga pinku no burausu-o kiteimasu ka?) Who-NOM pink blouse-ACC is wearing-Q → “Who is wearing a pink blouse?”	
1. Agent-HN: *a. Spanish:* ¿Quién lleva un sombrero marrón? Who wears a hat brown *b. Japanese:* 誰が茶色い帽子をかぶっていますか？ (Dare-ga chaairoi boushi-o kabutteimasu ka?) Who-NOM brown hat-ACC is wearing-Q → “Who is wearing a brown hat?”	2. Patient-HN: *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	

<u>1. Agent-HN:</u> *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	<u>2. Patient-HN:</u> *a. Spanish:* ¿Quién lleva unos pantalones verdes? Who wears some trousers green *b. Japanese:* 誰が緑色のズボンを履いていますか？ (Dare-ga midoriiro-no zubon-o haiteimasu ka?) Who-NOM green trousers-ACC is wearing-Q → “Who is wearing green trousers?”	

<u>Verb 2: Derribar / 倒す (Taosu) – “Knock down”</u>

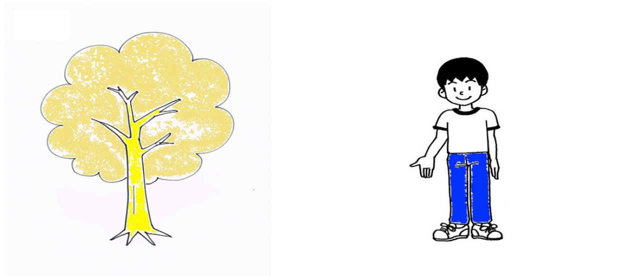
<u>1. Agent-HN:</u> *a. Spanish:* ¿Quién lleva un vestido negro? Who wears a dress black *b. Japanese:* 誰が黒いドレスを着ていますか？ (Dare-ga kuroi doresu-o kiteimasu ka?) Who-NOM black dress-ACC is wearing-Q → “Who is wearing a black dress?”	*2. Patient-HN:* *a. Spanish:* ¿Quién lleva una camisa marrón? Who wears a shirt brown *b. Japanese:* 誰が茶色いシャツを着ていますか？ (Dare-ga chairoi shatsu-o kiteimasu ka?) Who-NOM brown shirt-ACC is wearing-Q → “Who is wearing a brown shirt?”	

<table>
<tr>
<td>1. Agent-HN:
a. Spanish:
¿Quién lleva unos pantalones rojos?
Who wears some trousers red

b. Japanese:
誰が赤いズボンを履いていますか？
(Dare-ga akai zubon-o haiteimasu ka?)
Who-NOM red trousers-ACC is wearing-Q

→ “Who is wearing red trousers?”</td>
<td>2. Patient-HN:
a. Spanish:
¿Qué es verde?
What is green

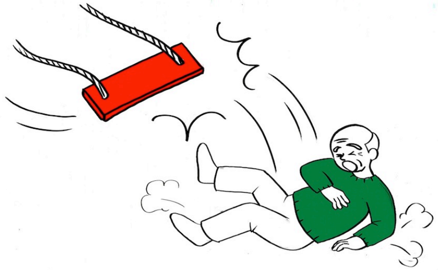
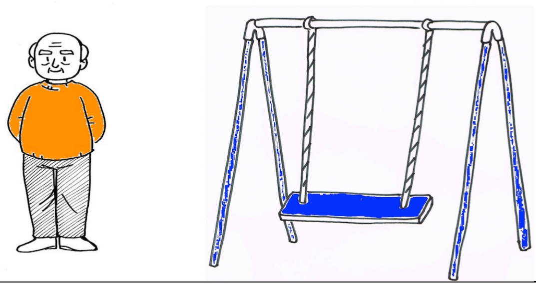
b. Japanese:
何が緑ですか？
(Nani-ga midori desu ka?)
What-NOM green is-Q

→ “What is green?”</td>
<td>
</td>
</tr>
<tr>
<td>1. Agent-HN:
a. Spanish:
¿Qué es rojo?
What is red

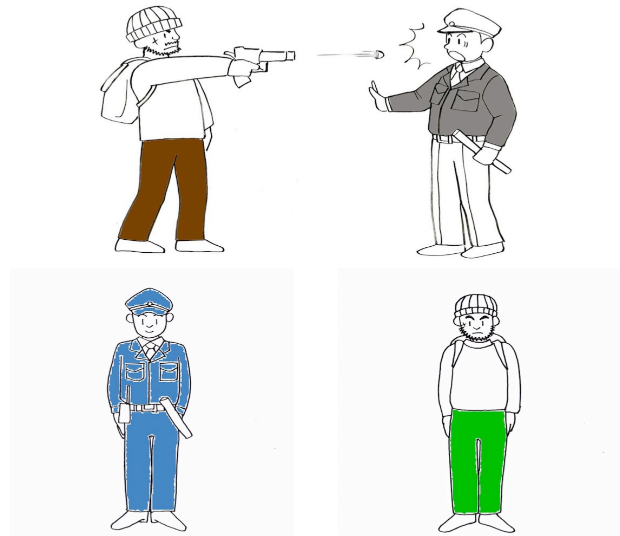
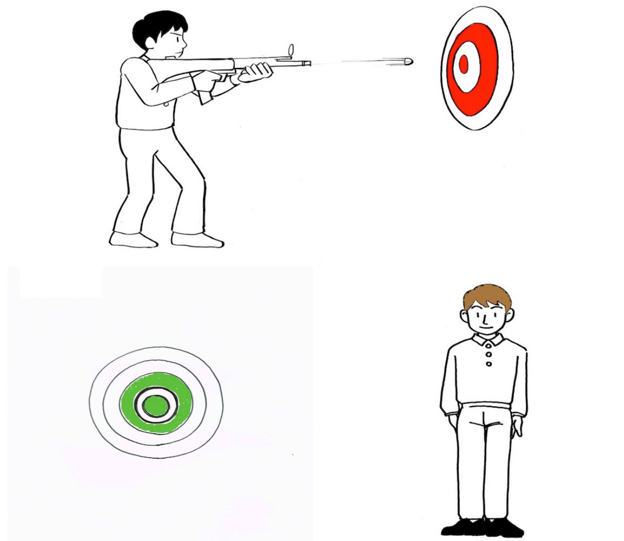
b. Japanese:
何が赤いですか？
(Nani-ga akai desu ka?)
What-NOM red is-Q

→ “What is red?”</td>
<td>2. Patient-HN:
a. Spanish:
¿Quién lleva un jersey verde?
Who wears a sweater green

b. Japanese:
誰が緑のゼーターを着ていますか？
(Dare-ga midori no seetaa-o kiteimasuka?)
Who-NOM green sweater-ACC is wearing-Q

→ “Who is wearing a green sweater?”</td>
<td>
</td>
</tr>
</table>

Verb 3: Disparar / 撃つ (Utsu) – “Shoot”

1. Agent-HN: *a. Spanish:* ¿Quién lleva unos pantalones marrones? Who wears some trousers brown *b. Japanese:* 誰が茶色いズボンを履いていますか？ (Dare-ga chairoi zubon-o haiteimasu ka?) Who-NOM brown trousers-ACC is wearing-Q → “Who is wearing brown trousers?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva una chaqueta gris? Who wears a jacket grey *b. Japanese:* 誰が灰色のジャケットを着ていますか？ (Dare-ga hairo-no jaketto-o kiteimasuka?) Who-NOM grey jacket-ACC is wearing-Q → “Who is wearing a grey jacket?”	
1. Agent-HN: *a. Spanish:* ¿Quién tiene el pelo negro? Who has the hair black *b. Japanese:* 誰が髪が黒いですか？ (Dare-ga kami-ga kuroi desu ka?) Who-NOM hair-NOM black is-Q → “Who has black hair?”	2. Patient-HN: *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	

1. Agent-HN:

a. Spanish:

¿Qué es gris?
What is grey

b. Japanese:

何が灰色ですか？
(Nani-ga haiiro desu ka?)
What-NOM grey is-Q

→ “What is grey?”

2. Patient-HN:

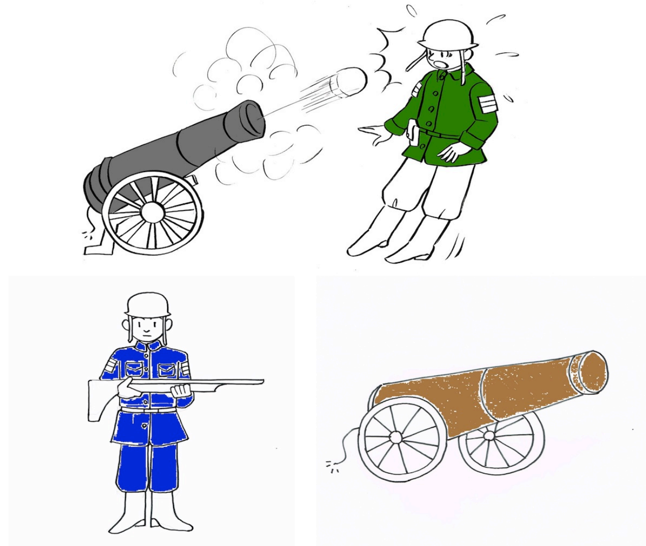
a. Spanish:

¿Quién lleva un uniforme verde?
Who wears an uniform gren

b. Japanese:

誰が緑色の制服を着ていますか？
(Dare-ga midoriiro-no seifuku-o kiteimasu ka?)
Who-NOM green uniform-ACC is wearing-Q

→ “Who is wearing a green uniform?”



Verb 4: Agarrar / 捉える (Toraeru) – “Grasp”

1. Agent-HN:

a. Spanish:

¿Quién lleva una chaqueta azul?
Who wears a jacket blue

b. Japanese:

誰が青いジャケットを着ていますか？
(Dare-ga aoi jaketto-o kiteimasu ka?)
Who-NOM blue jacket-ACC is wearing-Q

→ “Who is wearing a blue jacket?”

2. Patient-HN:

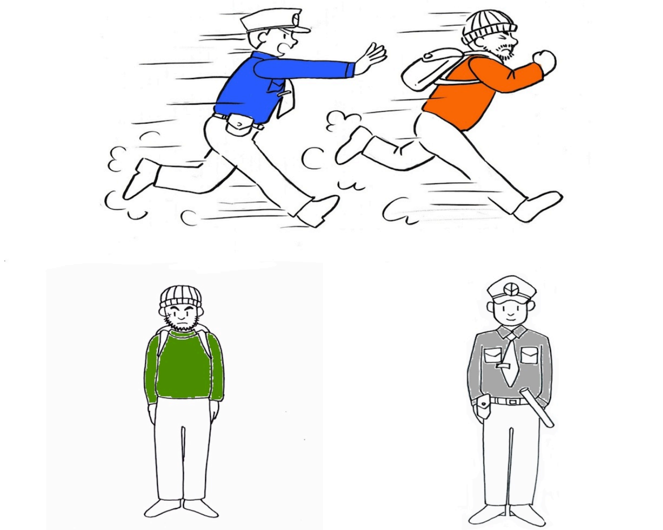
a. Spanish:

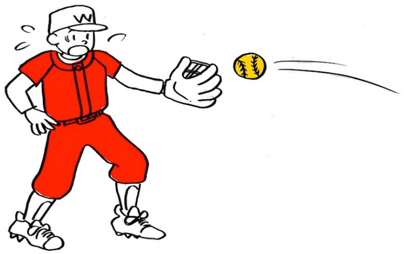
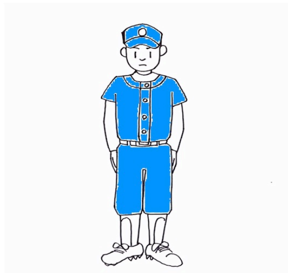
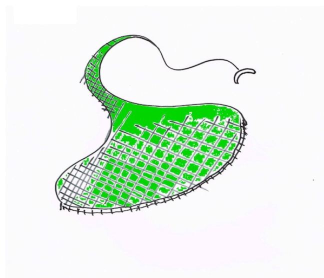
¿Quién lleva un jersey naranja?
Who wears a sweater orange

b. Japanese:

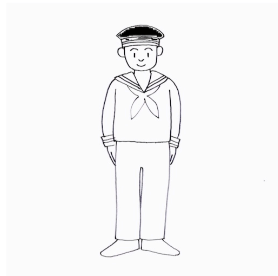
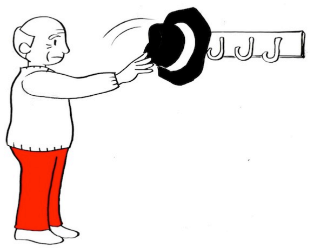
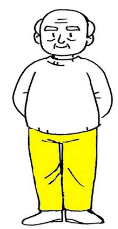
誰がオレンジセーターを着ていますか？
(Dare-ga orenji seetaa -o kiteimasu ka?)
Who-NOM orange sweater -ACC is wearing-Q

→ “Who is wearing an orange sweater ?



1. Agent-HN: *a. Spanish:* ¿Quién lleva un uniforme rojo? Who wears an uniform red *b. Japanese:* 誰が赤い制服を着ていますか？ (Dare-ga akai seifuku-o kiteimasu ka?) Who-NOM red uniform-ACC is wearing-Q → "Who is wearing a red uniform?"	2. Patient-HN: *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → "What is yellow?"	
1. Agent-HN: *a. Spanish:* ¿Qué es marrón? What is brown *b. Japanese:* 何が茶色いですか？ (Nani-ga chairoi desu ka?) What-NOM brown is-Q → "What is brown?"	2. Patient-HN: *a. Spanish:* ¿Quién lleva una camiseta negra? Who wears a T-shirt black *b. Japanese:* 誰が黒い T シャツを着ていますか？ (Dare-ga kuroi T-shatsu-o kiteimasu ka?) Who-NOM black T-shirt-ACC is wearing-Q → "Who is wearing a black T-shirt?"	

Verb 5: Enganchar / 引っ掛ける (Hikkakeru) – "Hook up"

1. Agent-HN: *a. Spanish:* ¿Quién lleva una chaqueta verde? Who wears a jacket green *b. Japanese:* 誰が緑色のジャケットを着ていますか？ (Dare-ga midoriiro-no jaketto-o kiteimasu ka?) Who-NOM green jacket-ACC is wearing-Q → "Who is wearing a green jacket?"	2. Patient-HN: *a. Spanish:* ¿Quién lleva un sombrero azul? Who wears a hat blue *b. Japanese:* 誰が青い帽子をかぶっていますか？ (Dare-ga aoi boushi-o kabutteimasu ka?) Who-NOM blue hat-ACC is wearing-Q → "Who is wearing a blue hat?"	  
1. Agent-HN: *a. Spanish:* ¿Quién lleva unos pantalones rojos? Who wears some trousers red *b. Japanese:* 誰が赤いズボンを履いていますか？ (Dare-ga akai zubon-o haiteimasu ka?) Who-NOM red trousers-ACC is wearing-Q → "Who is wearing red trousers?"	2. Patient-HN: *a. Spanish:* ¿Qué es negro? What is black *b. Japanese:* 何が黒いですか？ (Nani-ga kuroi desu ka?) What-NOM black is-Q → "What is black?"	  

1. Agent-HN:

a. Spanish:

¿Qué es amarillo?
What is yellow

b. Japanese:

何が黄色いですか？
(Nani-ga kiiroi desu ka?)
What-NOM yellow is-Q

→ "What is yellow?"

2. Patient-HN:

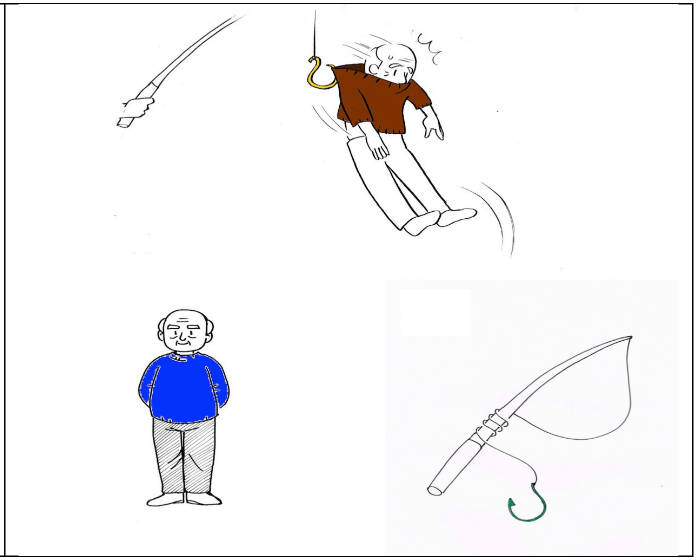
a. Spanish:

¿Quién lleva un jersey marrón?
Who wears a sweater brown

b. Japanese:

誰が茶色のゼーターを着ていますか？
(Dare-ga chairo no seetaa-o kiteimasu ka?)
Who-NOM brown sweater-ACC is wearing-Q

→ "Who is wearing a brown sweater?"



Verb 6: Golpear / 叩く (Tataku) – "Hit"

1. Agent-HN:

a. Spanish:

¿Quién lleva una blusa roja?
Who wears a blouse red

b. Japanese:

誰が赤いブラウスを着ていますか？
(Dare-ga akai burausu-o kiteimasu ka?)
Who-NOM red blouse-ACC is wearing-Q

→ "Who is wearing a red blouse?"

2. Patient-HN:

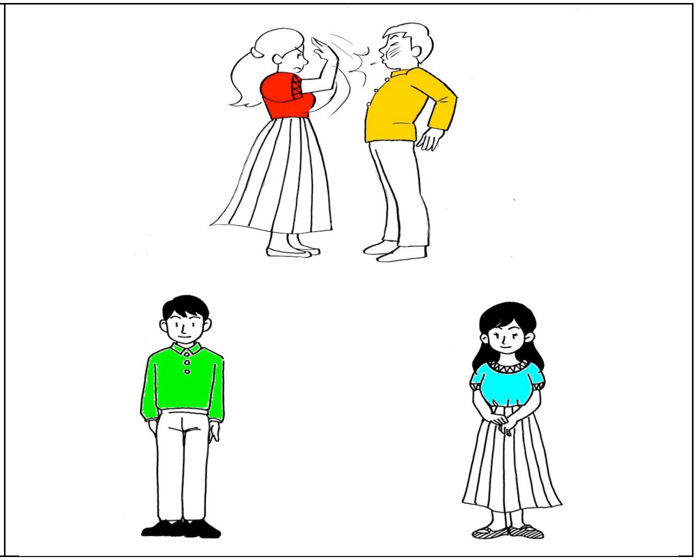
a. Spanish:

¿Quién lleva una camisa amarilla?
Who wears a shirt yellow

b. Japanese:

誰が黄色いシャツを着ていますか？
(Dare-ga kiiroi shatsu-o kiteimasu ka?)
Who-NOM yellow shirt-ACC is wearing-Q

→ "Who is wearing a yellow shirt?"

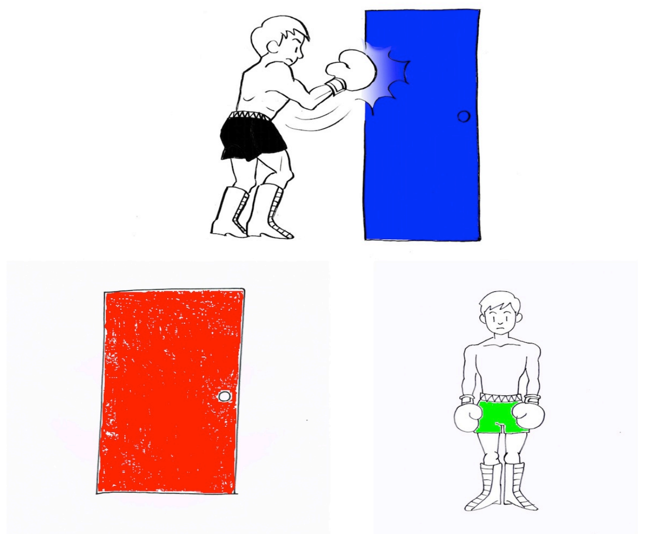


<table>
<tr>
<td><u>1. Agent-HN:</u>
a. Spanish:
¿Quién lleva unos pantalones negros?
Who wears some trousers black

b. Japanese:
誰が黒いズボンを履いていますか？
(Dare-ga kuroi zubon-o haiteimasu ka?)
Who-NOM black trousers-ACC is wearing-Q

→ “Who is wearing black trousers?”</td>
<td><u>2. Patient-HN:</u>
a. Spanish:
¿Qué es azul?
What is blue

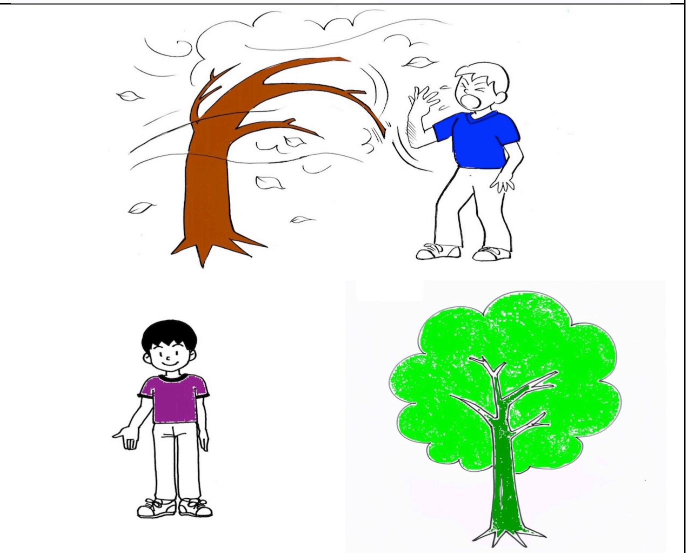
b. Japanese:
何が青いですか？
(Nani-ga aoi desu ka?)
What-NOM blue is-Q

→ “What is blue?”</td>
<td></td>
</tr>
<tr>
<td><u>1. Agent-HN:</u>
a. Spanish:
¿Qué es marrón?
What is brown

b. Japanese:
何が茶色いですか？
(Nani-ga chairoi desu ka?)
What-NOM brown is-Q

→ “What is brown?”</td>
<td><u>2. Patient-HN:</u>
a. Spanish:
¿Quién lleva una camiseta azul?
Who wears a T-shirt blue

b. Japanese:
誰が青いＴシャツを着ていますか？
(Dare-ga aoi T-shatsu-o kiteimasu ka?)
Who-NOM blue T-shirt-ACC is wearing-Q

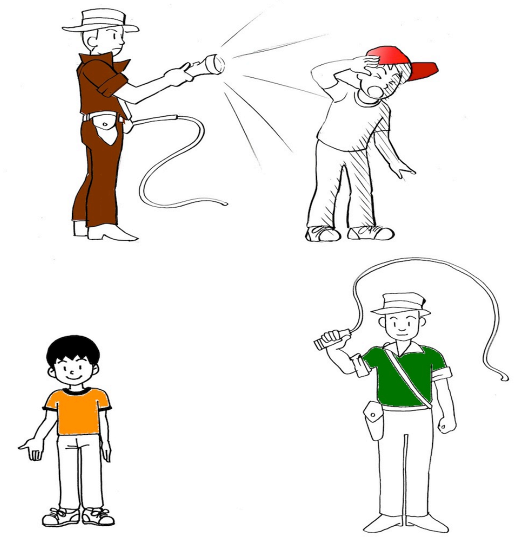
→ “Who is wearing a blue T-shirt?”</td>
<td></td>
</tr>
</table>

<table>
<tr>
<td>1. Agent-HN:
a. Spanish:
¿Quién lleva una chaqueta marrón?
Who wears a jacket brown

b. Japanese:
誰が茶色のジャケットを着ていますか？
(Dare-ga chaiiro-no jaketto-o kiteimasu ka?)
Who-NOM brown jacket-ACC is wearing-Q

→ “Who is wearing a brown jacket?”</td>
<td>2. Patient-HN:
a. Spanish:
¿Quién lleva una gorra roja?
Who wears a cap red

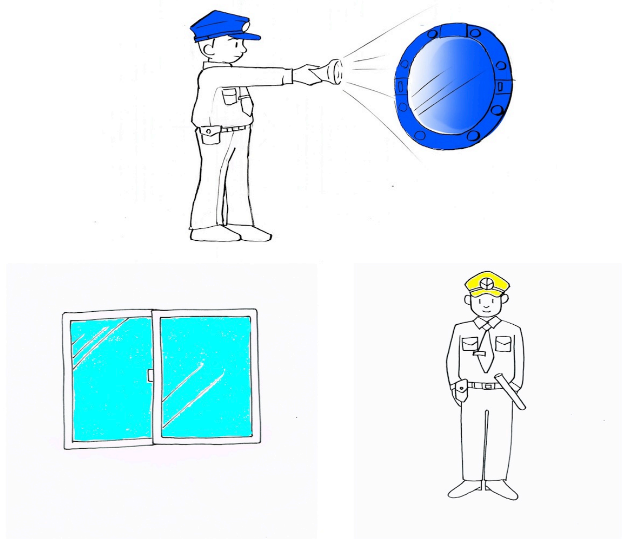
b. Japanese:
誰が赤いキャップをかぶっていますか？
(Dare-ga akai kyappu-o kabutteimasu ka?)
Who-NOM red cap-ACC is wearing-Q

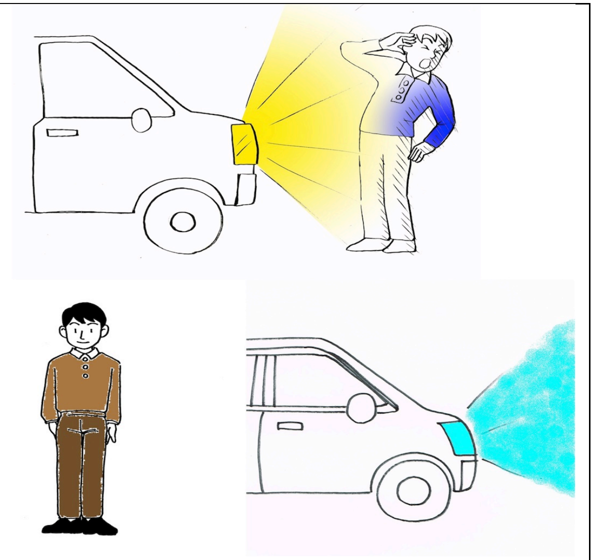
→ “Who is wearing a red cap?”</td>
<td></td>
</tr>
<tr>
<td>1. Agent-HN:
a. Spanish:
¿Quién lleva un sombrero azul?
Who wears a hat blue

b. Japanese:
誰が青い帽子をかぶっていますか？
(Dare-ga aoi boushi-o kabutteimasu ka?)
Who-NOM blue hat-ACC is wearing-Q

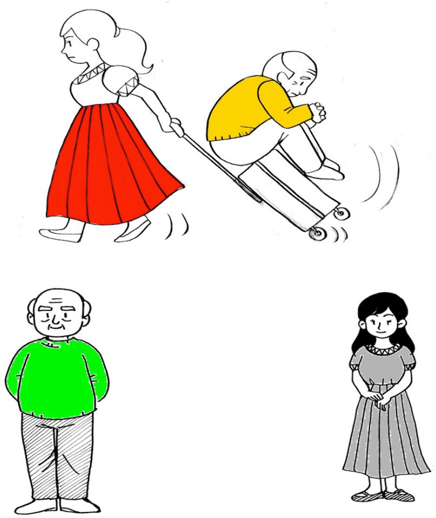
→ “Who is wearing a blue hat?”</td>
<td>2. Patient-HN:
a. Spanish:
¿Qué es redondo?
What is rounded

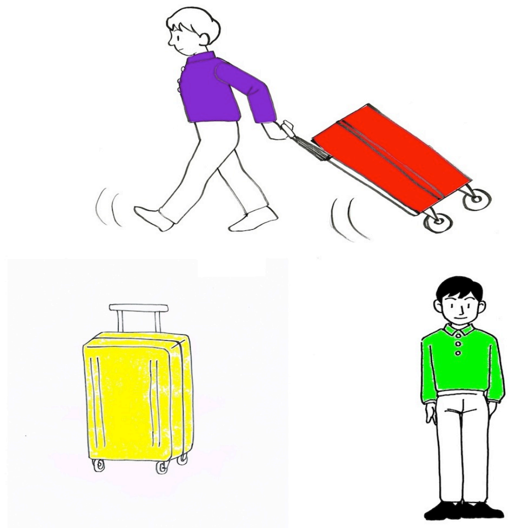
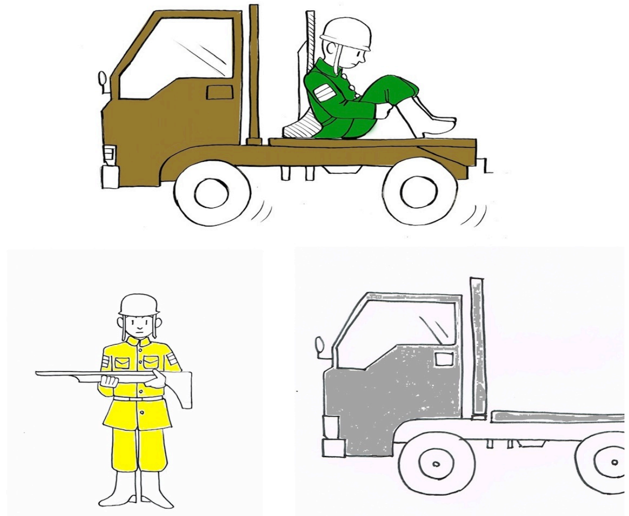
b. Japanese:
何が丸いですか？
(Nani-ga marui desu ka?)
What-NOM rounded is-Q

→ “What is rounded?”</td>
<td></td>
</tr>
</table>

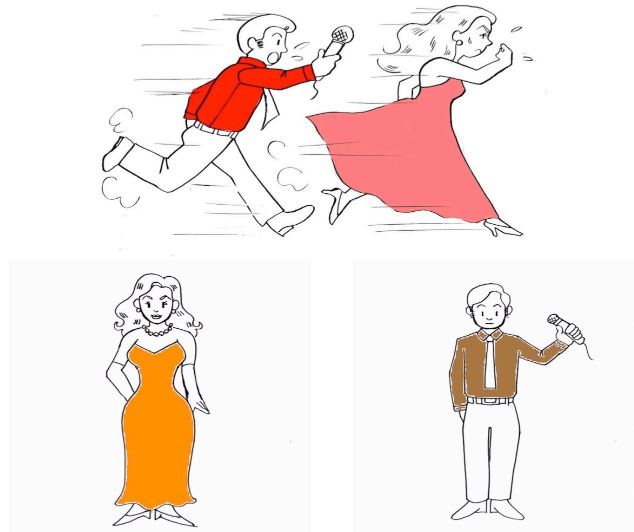
1. Agent-HN: *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva una camisa azul? Who wears a shirt blue *b. Japanese:* 誰が青いシャツを着ていますか？ (Dare-ga aoi shatsu-o kiteimasu ka?) Who-NOM blue shirt-ACC is wearing-Q → “Who is wearing a blue shirt?”	

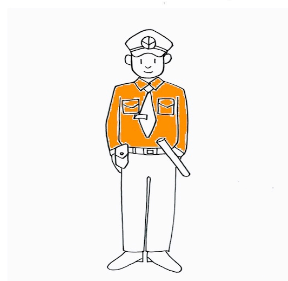
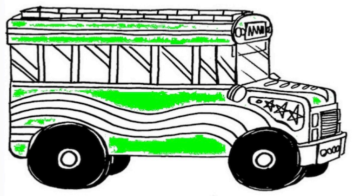
Verb 8: Llevar / 運ぶ (Hakobu) – “Carry”

1. Agent-HN: *a. Spanish:* ¿Quién lleva una falda roja? Who wears a skirt red *b. Japanese:* 誰が赤いスカートを履いていますか？ (Dare-ga akai sukaato-o haiteimasu ka?) Who-NOM red skirt-ACC is wearing-Q → “Who is wearing a red skirt?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva un jersey amarillo? Who wears a sweater yellow *b. Japanese:* 誰が黄色いゼーターを着ていますか？ (Dare-ga kiiroi seetaa-o kiteimasu ka?) Who-NOM yellow sweater-ACC is wearing-Q → “Who is wearing a yellow sweater?”	

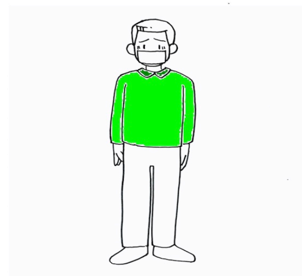
<u>1. Agent-HN:</u> *a. Spanish:* ¿Quién lleva una camisa morada? Who wears a shirt purple *b. Japanese:* 誰が紫のシャツを着ていますか？ (Dare-ga murasaki no shatsu-o kiteimasu ka?) Who-NOM purple shirt-ACC is wearing-Q → “Who is wearing a purple shirt?”	<u>2. Patient-HN:</u> *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	
<u>1. Agent-HN:</u> *a. Spanish:* ¿Qué es marrón? What is brown *b. Japanese:* 何が茶色いですか？ (Nani-ga chairoi desu ka?) What-NOM brown is-Q → “What is brown?”	<u>2. Patient-HN:</u> *a. Spanish:* ¿Quién lleva un uniforme verde? Who wears an uniform gren *b. Japanese:* 誰が緑色の制服を着ていますか？ (Dare-ga midoriiro-no seifuku-o kiteimasu ka?) Who-NOM green uniform-ACC is wearing-Q → “Who is wearing a green uniform?”	

Verb 9: Perseguir / 追い掛ける (Oikakeru) – “Chase”

1. Agent-HN: *a. Spanish:* ¿Quién lleva una camisa roja? Who wears a shirt red *b. Japanese:* 誰が赤いシャツを着ていますか？ (Dare-ga akai shatsu-o kiteimasu ka?) Who-NOM red shirt-ACC is wearing-Q → “Who is wearing a red shirt?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva un vestido rosa? Who wears a dress pink *b. Japanese:* 誰がピンクのドレスを着ていますか？ (Dare-ga pinku no doresu-o kiteimasu ka?) Who-NOM pink dress-ACC is wearing-Q → “Who is wearing a pink dress?”	
1. Agent-HN: *a. Spanish:* ¿Quién lleva una falda morada? Who wears a skirt purple *b. Japanese:* 誰が紫のスカートを履いていますか？ (Dare-ga murasaki no sukaato-o haiteimasu ka?) Who-NOM purple skirt-ACC is wearing-Q → “Who is wearing a purple skirt?”	2. Patient-HN: *a. Spanish:* ¿Qué es negro? What is black *b. Japanese:* 何が黒いですか？ (Nani-ga kuroi desu ka?) What-NOM black is-Q → “What is black?”	

1. Agent-HN: *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva una chaqueta amarilla? Who wears a jacket grey *b. Japanese:* 誰が黄色いジャケットを着ていますか？ (Dare-ga kiiroi jaketto-o kiteimasu ka?) Who-NOM yellow jacket-ACC is wearing-Q → “Who is wearing a yellow jacket?”	  

Verb 10: Pinchar / 刺す (Sasu) – “Poke”

1. Agent-HN: *a. Spanish:* ¿Quién lleva una bata azul? Who wears a coat blue *b. Japanese:* 誰が青い白衣を着ていますか？ (Dare-ga aoi hakui-o kiteimasu ka?) Who-NOM blue coat-ACC is wearing-Q → “Who is wearing a blue coat?”	2. Patient-HN: *a. Spanish:* ¿Quién lleva una camisa morada? Who wears a shirt purple *b. Japanese:* 誰が紫のシャツを着ていますか？ (Dare-ga murasaki no shatsu-o kiteimasu ka?) Who-NOM purple shirt-ACC is wearing-Q → “Who is wearing a purple shirt?”	  

1. Agent-HN:

a. Spanish:

¿Quién lleva un mono azul?
Who wears a jumpsuit blue

b. Japanese:

誰が青いつなぎを着ていますか？
(Dare-ga aoi tsunagi-o kiteimasu ka?)
Who-NOM blue jumpsuit-ACC is wearing-Q

→ “Who is wearing a blue jumpsuit?”

2. Patient-HN:

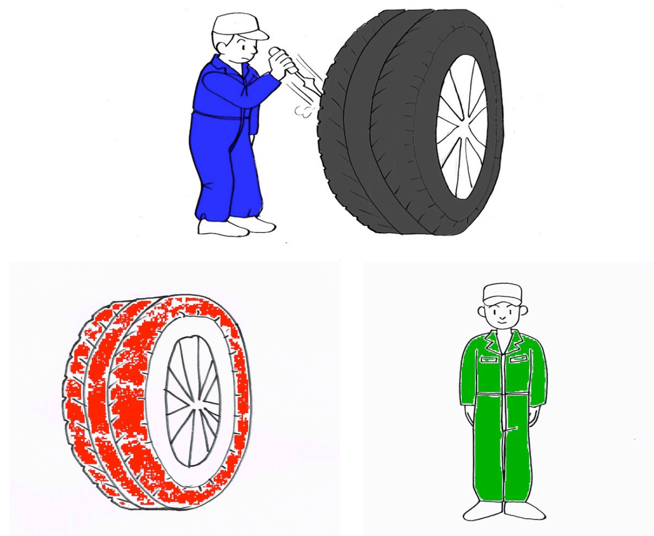
a. Spanish:

¿Qué es negro?
What is black

b. Japanese:

何が黒いですか？
(Nani-ga kuroi desu ka?)
What-NOM black is-Q

→ “What is black?”



1. Agent-HN:

a. Spanish:

¿Qué es verde?
What is green

b. Japanese:

何が緑ですか？
(Nani-ga midori desu ka?)
What-NOM green is-Q

→ “What is green?”

2. Patient-HN:

a. Spanish:

¿Quién lleva una blusa azul?
Who wears a blouse blue

b. Japanese:

誰が青いブラウスを着ていますか？
(Dare-ga aoi burausu-o kiteimasu ka?)
Who-NOM blue blouse-ACC is wearing-Q

→ “Who is wearing a blue blouse?”

Appendix 2: List of filler items

Fillers items include a total of 30, hand-drawn pictures presented with questions. Each of the items was presented twice during the experiment (once per block) with a question referred to either the upper or the lower element. Questions in both cases are presented in the list below.

There were three different types of fillers:

4. Fillers that represent intransitive actions in both Spanish and Japanese
5. Fillers that represent intransitive actions in Japanese but stative events in Spanish.
6. Fillers that represent size contrast

The two first types of items are presented mixed in the list. In them, a verb is provided for each of the pictures (presented in parenthesis before each question in the list below). The first type presented another type of sentences. The question refers again to a feature of one of the elements, but in this case no verb is provided. In its place, the word "size" in each of the targeted languages, is provided. Participants were instructed during the trials that in these cases, they have to answer with a size contrast (stating whether the element is big or small).

Filler items were not counterbalanced between participants and were presented as they are listed below (in a randomized order). However, we ensured that each participant saw a mix of items with the target element at the top and at the bottom.

List of fillers with provided questions in Japanese and Spanish, and provided verb.

Items with intrintransitive or stative verb

1. Top left: Sonrojarse / 照れる (“Blush”)

a. Spanish:

¿Quién lleva unos pantalones amarillos?
Who wears some trousers yellow

b. Japanese:

誰が黄色いズボンを履いていますか？
(Dare-ga kiiroi zubon-o haiteimasu ka?)
Who-NOM yellow trousers-ACC is wearing-Q

→ “Who is wearing yellow trousers?”

2. Bottom right: Reír / 笑う (“Laugh”)

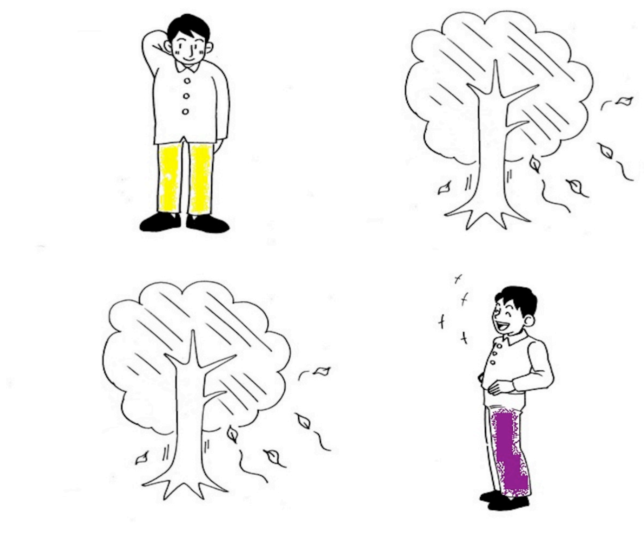
a. Spanish:

¿Quién lleva unos pantalones morados?
Who wears some trousers purple

b. Japanese:

誰が紫のズボンを履いていますか？
(Dare-ga murasaki-no zubon-o haiteimasu ka?)
Who-NOM purple trousers-ACC is wearing-Q

→ “Who is wearing purple trousers?”



1. Top left: Enfadarse / 怒る (“Be mad”)

a. Spanish:

¿Quién lleva una camiseta roja?
Who wears a T-shirt red

b. Japanese:

誰が赤いＴシャツを着ていますか？
(Dare-ga akai T-shatsu-o kiteimasu ka?)
Who-NOM red T-shirt-ACC is wearing-Q

→ “Who is wearing a red T-shirt?”

2. Bottom right: Llorar / 泣く (“Cry”)

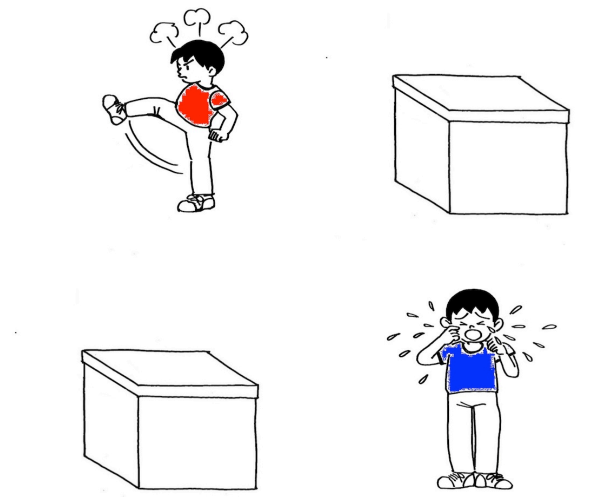
a. Spanish:

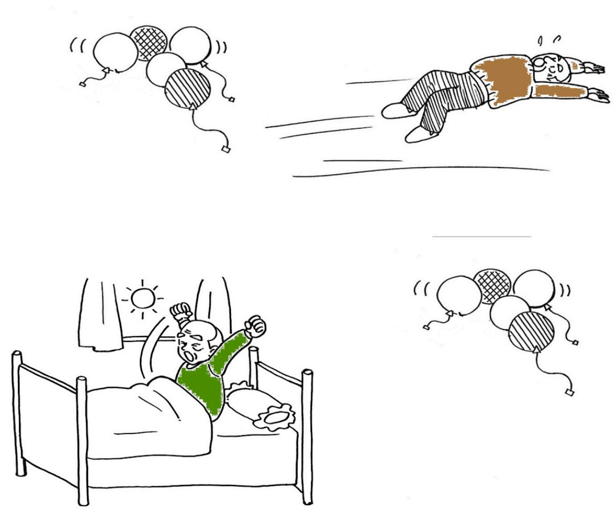
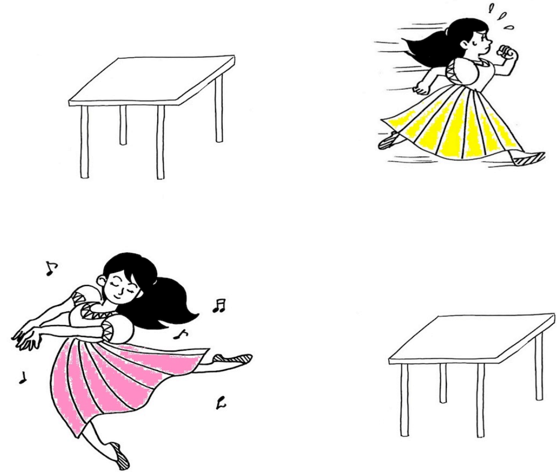
¿Quién lleva una camiseta azul?
Who wears a T-shirt blue

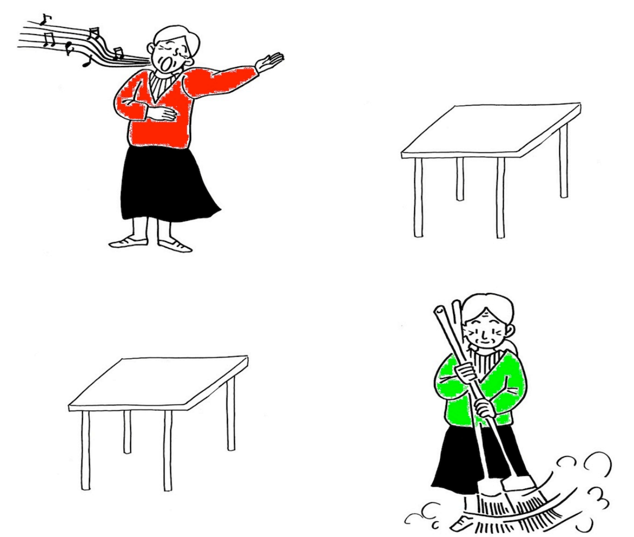
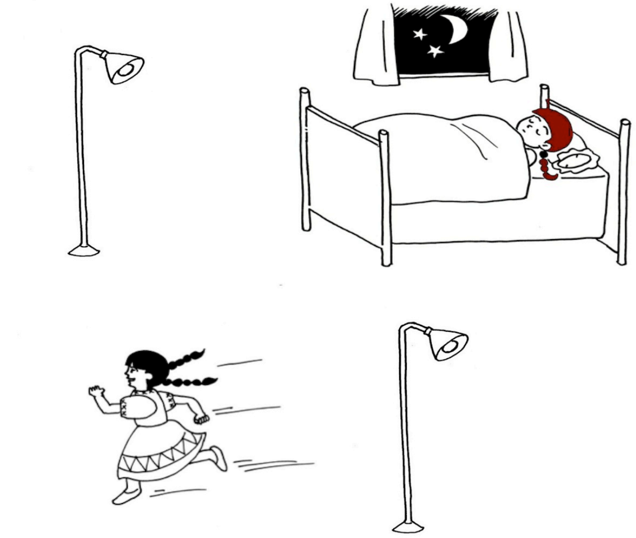
b. Japanese:

誰が青いＴシャツを着ていますか？
(Dare-ga aoi T-shatsu-o kiteimasu ka?)
Who-NOM blue T-shirt-ACC is wearing-Q

→ “Who is wearing a blue T-shirt?”



1. Top right: Volar / 飛ぶ (“Fly”) *a. Spanish:* ¿Quién lleva un jersey marrón? Who wears a sweater brown *b. Japanese:* 誰が茶色のセーターを着ていますか？ (Dare-ga chairo-no seetaa-o kiteimasu ka?) Who-NOM brown sweater-ACC is wearing-Q → “Who is wearing a brown sweater?”	2. Bottom left: Despertarse / 起きる (“Wake up”) *a. Spanish:* ¿Quién lleva un jersey verde? Who wears a sweater green *b. Japanese:* 誰が緑のセーターを着ていますか？ (Dare-ga midori-no seetaa-o kiteimasu ka?) Who-NOM green sweater-ACC is wearing-Q → “Who is wearing a green sweater?”	
1. Top right: Correr / 走る (“Run”) *a. Spanish:* ¿Quién lleva una falda amarilla? Who wears a skirt yellow *b. Japanese:* 誰が黄色いスカートを履いていますか？ (Dare-ga kiiroi sukaato-o haiteimasu ka?) Who-NOM yellow skirt-ACC is wearing-Q → “Who is wearing a yellow skirt?”	2. Bottom left: Bailar / 踊る (“Dance”) *a. Spanish:* ¿Quién lleva una falda rosa? Who wears a skirt pink *b. Japanese:* 誰がピンクのスカートを履いていますか？ (Dare-ga pinku-no sukaato-o haiteimasu ka?) Who-NOM pink skirt-ACC is wearing-Q → “Who is wearing a pink skirt?”	

1. Top left: Cantar / 歌う (“Sing”) *a. Spanish:* ¿Quién lleva un jersey rojo? Who wears a sweater red *b. Japanese:* 誰が赤いセーターを着ていますか？ (Dare-ga akai seetaa-o kiteimasu ka?) Who-NOM red sweater-ACC is wearing-Q → “Who is wearing a red sweater?”	2. Bottom right: Barrer / 掃く (“Sweep”) *a. Spanish:* ¿Quién lleva un jersey verde? Who wears a sweater green *b. Japanese:* 誰が緑のゼーターを着ていますか？ (Dare-ga midori no seetaa-o kiteimasuka?) Who-NOM green sweater-ACC is wearing-Q → “Who is wearing a green sweater?”	
1. Top right: Dormir / 寝る (“Sleep”) *a. Spanish:* ¿Quién tiene el pelo rojo? Who has the hair red *b. Japanese:* 誰が髪が赤いですか？ (Dare-ga kami-ga akai ka?) Who-NOM hair-NOM red is-Q → “Who has red hair?”	2. Bottom left: Correr / 走る (“Run”) *a. Spanish:* ¿Quién tiene el pelo negro? Who has the hair black *b. Japanese:* 誰が髪が黒いですか？ (Dare-ga kami-ga kuroi ka?) Who-NOM hair-NOM black is-Q → “Who has black hair?”	

1. Top right: Caer / 落ちる (“Fall”)

a. Spanish:

¿Qué es marrón?
What is brown

b. Japanese:

何が茶色いですか？
(Nani-ga chairoi desu ka?)
What-NOM brown is-Q

→ “What is brown?”

2. Bottom left: Volar / 飛ぶ (“Fly”)

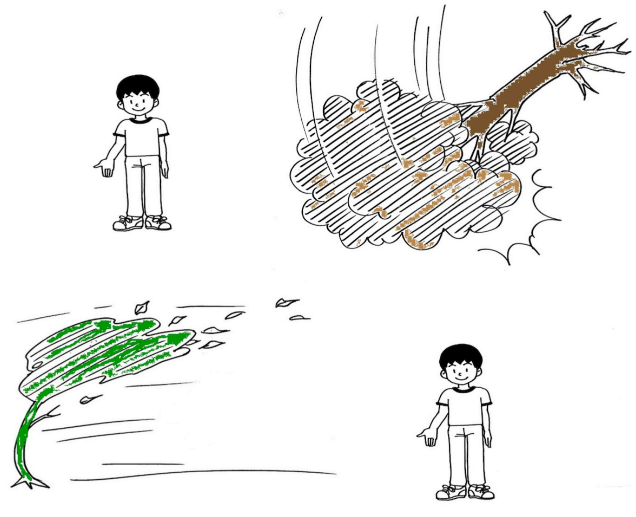
a. Spanish:

¿Qué es verde?
What is green

b. Japanese:

何が緑ですか？
(Nani-ga midori desu ka?)
What-NOM green is-Q

→ “What is green?”



1. Top right: Dormir / 寝る (“Sleep”)

a. Spanish:

¿Quién tiene el pelo marrón?
Who has the hair brown

b. Japanese:

誰が髪が茶色いですか？
(Dare-ga kami-ga chairoi ka?)
Who-NOM hair-NOM brown is-Q

→ “Who has brown hair?”

2. Bottom left: Despertarse / 起きる (“Wake up”)

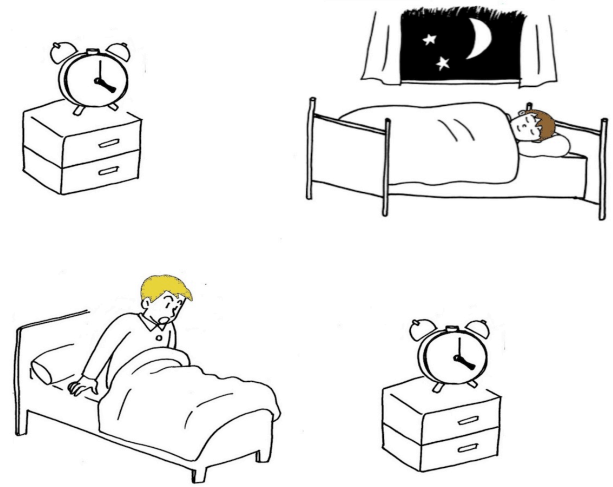
a. Spanish:

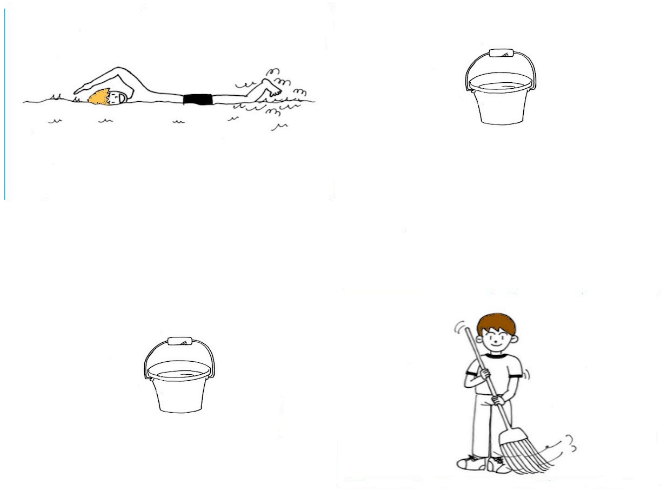
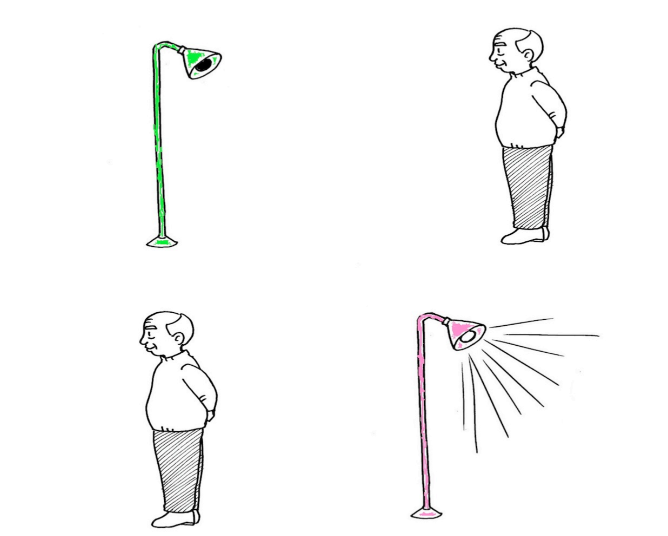
¿Quién tiene el pelo amarillo?
Who has the hair yellow

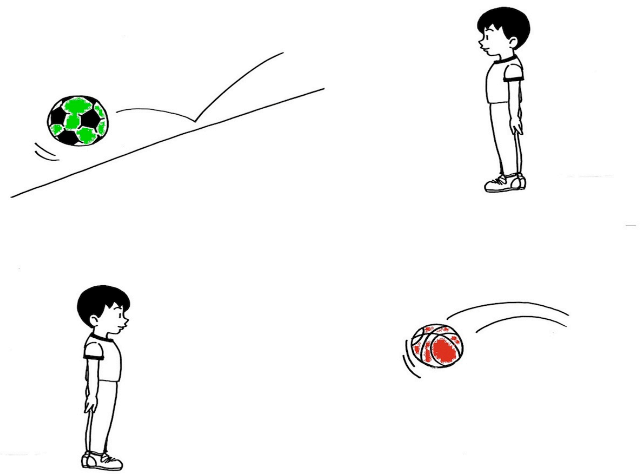
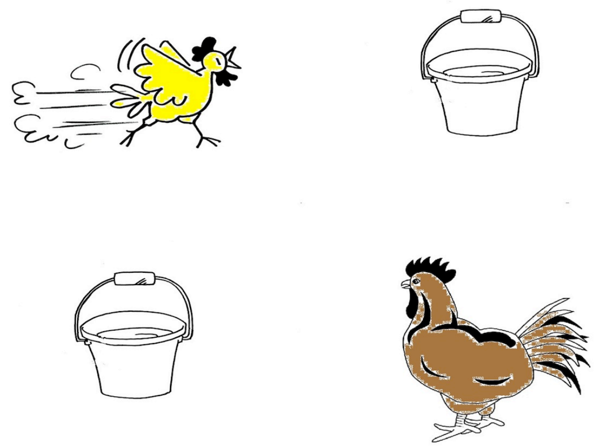
b. Japanese:

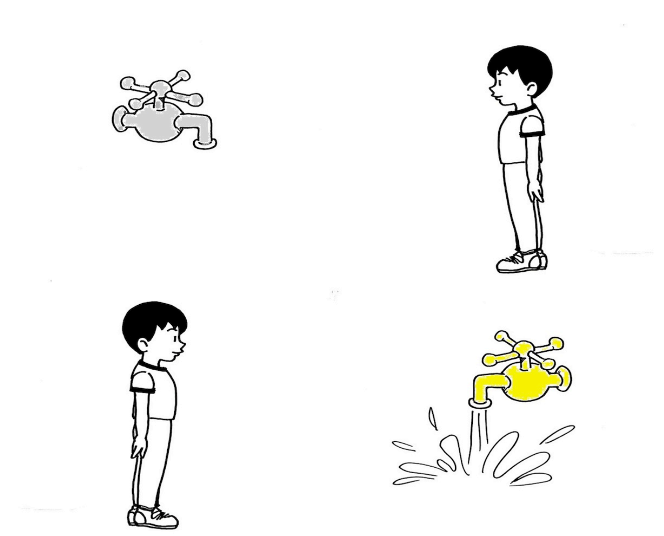
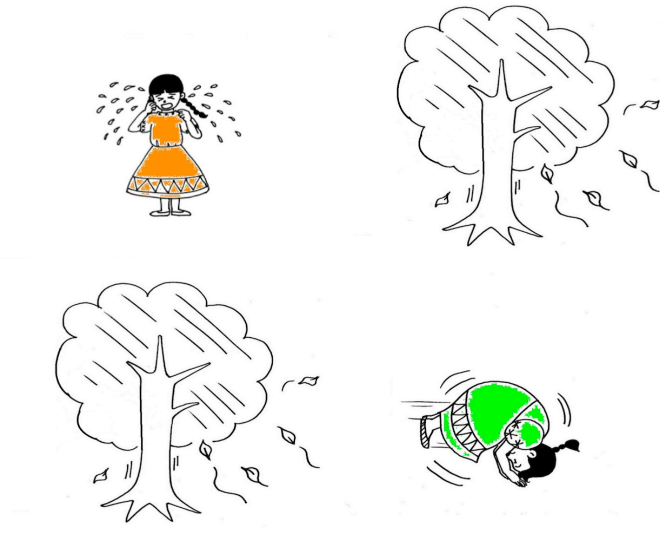
誰が髪が黄色いですか？
(Dare-ga kami-ga kiiroi ka?)
Who-NOM hair-NOM yellow is-Q

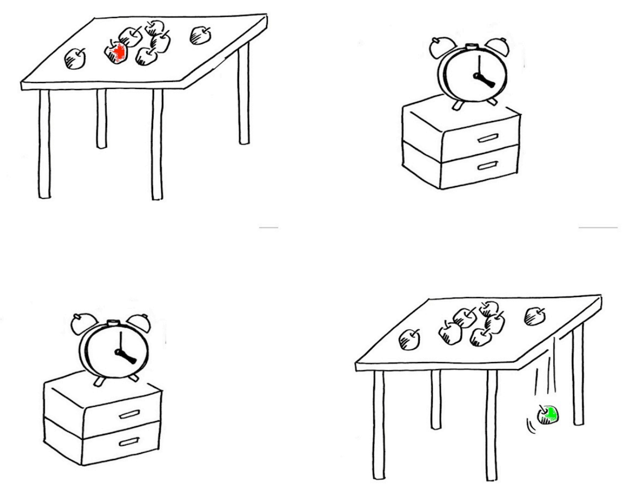
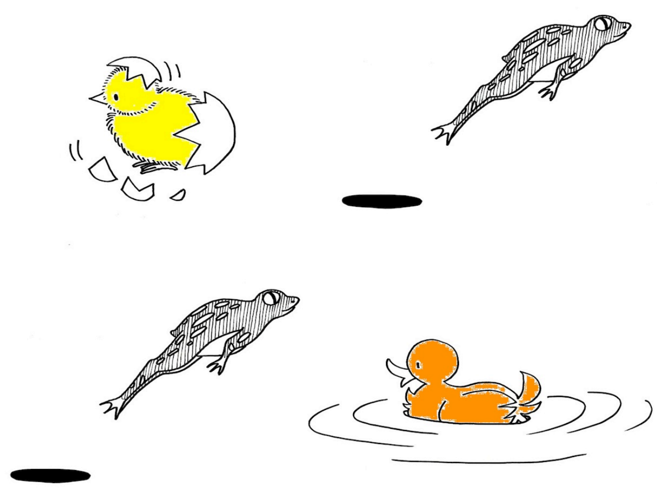
→ “Who has yellow hair?”

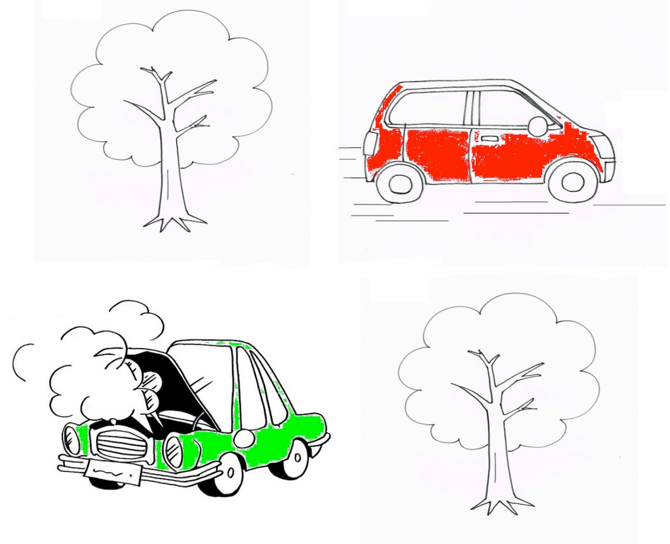


1. Top left: Nadar / 泳ぐ (“Swim”) *a. Spanish:* ¿Quién tiene el pelo amarillo? Who has the hair yellow *b. Japanese:* 誰が髪が黄色いですか？ (Dare-ga kami-ga kiiroi ka?) Who-NOM hair-NOM yellow is-Q → “Who has yellow hair?”	2. Bottom right: Barrer / 掃く (“Sweep”) *a. Spanish:* ¿Quién tiene el pelo marrón? Who has the hair brown *b. Japanese:* 誰が髪が茶色いですか？ (Dare-ga kami-ga chairoi ka?) Who-NOM hair-NOM brown is-Q → “Who has brown hair?”	
1. Top left: Apagar / 消す (“Switch off”) *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	2. Bottom right: Encender / 付ける (“Turn on”) *a. Spanish:* ¿Qué es rosa? What is pink *b. Japanese:* 何がピンクですか？ (Nani-ga pinku desu ka?) What-NOM pink is-Q → “What is pink?”	

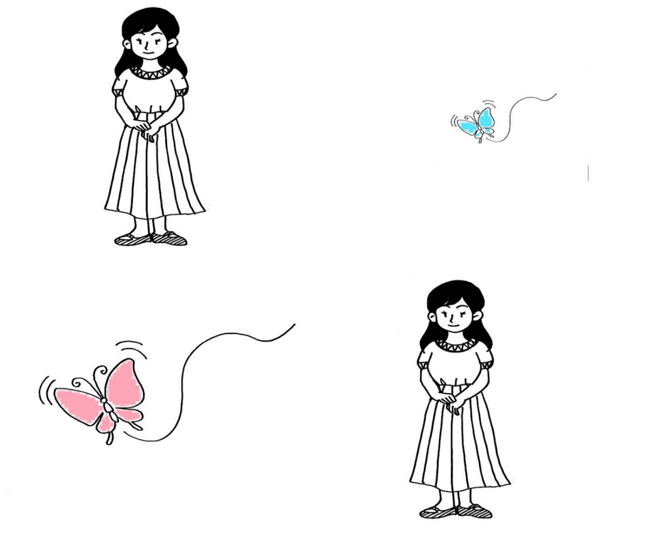
1. Top left: Rodar / 転ぶ (“Roll”) *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	2. Bottom right: Volar / 飛ぶ (“Fly”) *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	
1. Top left: Correr / 走る (“Run”) *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	2. Bottom right: Parar / 止まる (“Stop”) *a. Spanish:* ¿Qué es marrón? What is brown *b. Japanese:* 何が茶色いですか？ (Nani-ga chairoi desu ka?) What-NOM brown is-Q → “What is brown?”	

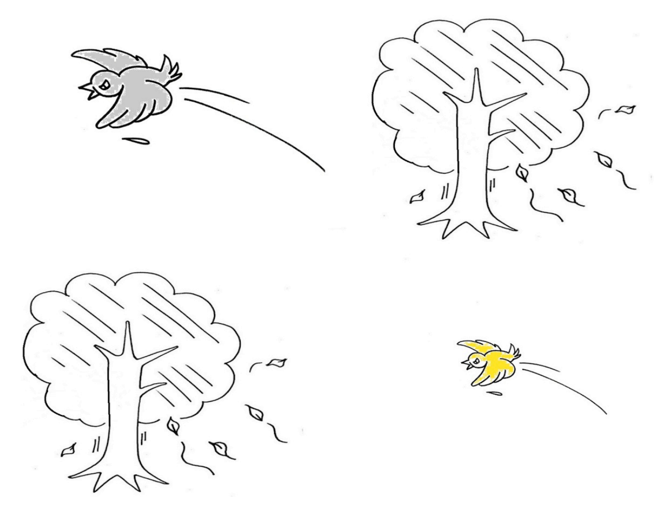
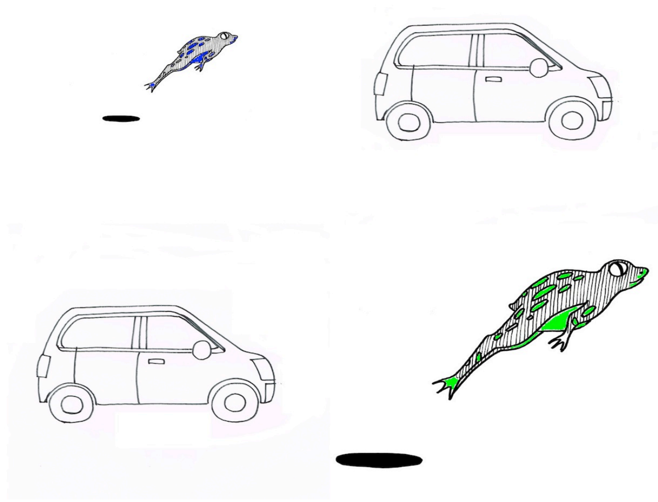
<u>1. Top left: Cerrar / 閉まる (“Close”)</u> *a. Spanish:* ¿Qué es gris? What is grey *b. Japanese:* 何が灰色ですか？ (Nani-ga haiiro desu ka?) What-NOM grey is-Q → “What is grey?”	<u>2. Bottom right: Abrir / 開く (“Open”)</u> *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	
<u>1. Top left: Llorar / 泣く (“Cry”)</u> *a. Spanish:* ¿Quién lleva un vestido naranja? Who wears a dress orange *b. Japanese:* 誰がオレンジのドレスを着ていますか？ (Dare-ga orenji no doresu-o kiteimasu ka?) Who-NOM orange dress-ACC is wearing-Q → “Who is wearing an orange dress?”	<u>2. Bottom right: Rodar / 転ぶ (¨Roll¨)</u> *a. Spanish:* ¿Quién lleva un vestido verde? Who wears a dress green *b. Japanese:* 誰が緑のドレスを着ていますか？ (Dare-ga midori no doresu-o kiteimasu ka?) Who-NOM green dress-ACC is wearing-Q → “Who is wearing a green dress?”	

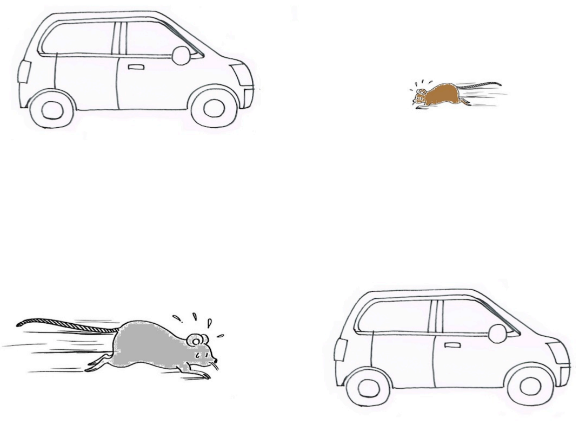
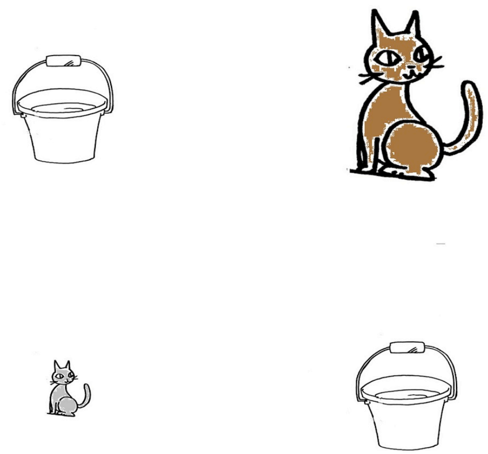
1. Top left: Estar / ある (“Stay”) *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	2. Bottom right: Caer / 落ちる (“Fall”) *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	
1. Top left: Nacer / 生まれる (“Born”) *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	1. Bottom right: Nadar / 泳ぐ (“Swim”) *a. Spanish:* ¿Qué es naranja? What is orange *b. Japanese:* 何がオレンジですか？ (Nani-ga orenji desu ka?) What-NOM orange is-Q → “What is orange?”	

1. Top right: Circular / 走る (“Circulate”) *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	2. Bottom left: Romper / 壊れる (“Break”) *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	
1. Top right: Enfadar / 怒る (“Be mad”) *a. Spanish:* ¿Quién lleva un jersey morado? Who wears a sweater purple *b. Japanese:* 誰が紫のセーターを着ていますか？ (Dare-ga murasaki-no seetaa-o kiteimasu ka?) Who-NOM purple sweater-ACC is wearing-Q → “Who is wearing a purple sweater?”	1. Bottom left: Sentarse / 座る (“Sit down”) *a. Spanish:* ¿Quién lleva un jersey verde? Who wears a sweater green *b. Japanese:* 誰が緑のセーターを着ていますか？ (Dare-ga midori-no seetaa-o kiteimasu ka?) Who-NOM green sweater-ACC is wearing-Q → “Who is wearing a green sweater?”	

Items with size contrast: provided word – Tamaño / サイズ (“Size”)

1. Top right: *a. Spanish:* ¿Qué es azul? What is blue *b. Japanese:* 何が青いですか？ (Nani-ga aoi desu ka?) What-NOM red is-Q → “What is blue?”	2. Bottom left: *a. Spanish:* ¿Qué es rosa? What is blue *b. Japanese:* 何がピンクですか？ (Nani-ga pinku desu ka?) What-NOM pink is-Q → “What is pink?”	
1. Top left: *a. Spanish:* ¿Qué es marrón? What is marrón *b. Japanese:* 何が茶色いですか？ (Nani-ga chairoi desu ka?) What-NOM brown is-Q → “What is brown?”	2. Bottom right: *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → “What is yellow?”	

1. Top left: *a. Spanish:* ¿Qué es gris? What is grey *b. Japanese:* 何が灰色ですか？ (Nani-ga haiiro desu ka?) What-NOM grey is-Q → "What is grey?"	2. Bottom right: *a. Spanish:* ¿Qué es amarillo? What is yellow *b. Japanese:* 何が黄色いですか？ (Nani-ga kiiroi desu ka?) What-NOM yellow is-Q → "What is yellow?"	
1. Top left: *a. Spanish:* ¿Qué es azul? What is blue *b. Japanese:* 何が青いですか？ (Nani-ga aoi desu ka?) What-NOM blue is-Q → "What is blue?"	2. Bottom right: *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → "What is green?"	

1. Top right: *a. Spanish:* ¿Qué es marrón? What is brown *b. Japanese:* 何が茶色ですか？ (Nani-ga chairo desu ka?) What-NOM brown is-Q → “What is brown?”	2. Bottom left: *a. Spanish:* ¿Qué es gris? What is grey *b. Japanese:* 何が灰色ですか？ (Nani-ga haiiro desu ka?) What-NOM grey is-Q → “What is grey?”	
1. Top right: *a. Spanish:* ¿Qué es marrón? What is brown *b. Japanese:* 何が茶色ですか？ (Nani-ga chairo desu ka?) What-NOM brown is-Q → “What is brown?”	2. Bottom left: *a. Spanish:* ¿Qué es gris? What is grey *b. Japanese:* 何が灰色ですか？ (Nani-ga haiiro desu ka?) What-NOM grey is-Q → “What is grey?”	

1. Top left:

a. Spanish:

¿Qué es morado?

What is purple

b. Japanese:

何が紫ですか？

(Nani-ga murasaki desu ka?)

What-NOM purple is-Q

→ “What is purple?”

2. Bottom right:

a. Spanish:

¿Qué es naranja?

What is orange

b. Japanese:

何がオレンジですか？

(Nani-ga orenji desu ka?)

What-NOM orange is-Q

→ “What is orange?”

1. Top left:

a. Spanish:

¿Qué es amarillo?

What is yellow

b. Japanese:

何が黄色いですか？

(Nani-ga kiiroi desu ka?)

What-NOM yellow is-Q

→ “What is yellow?”

2. Bottom right:

a. Spanish:

¿Qué es rojo?

What is red

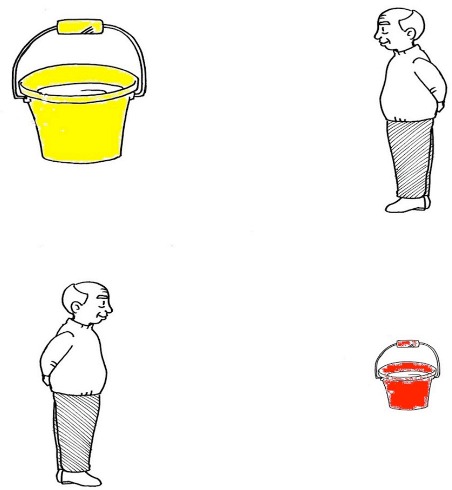
b. Japanese:

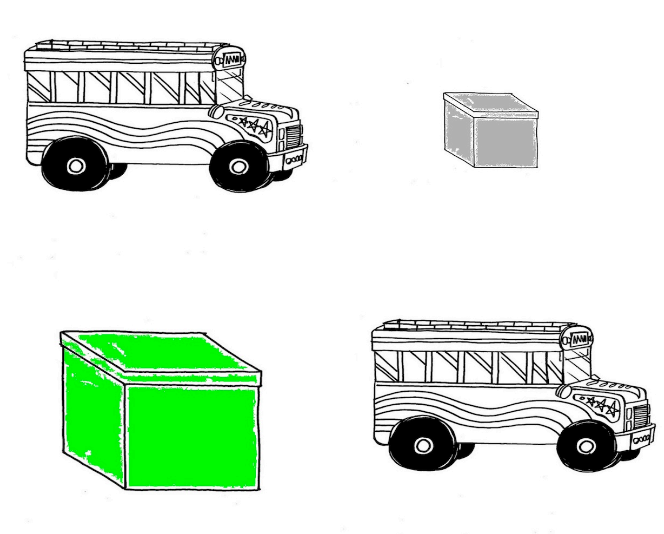
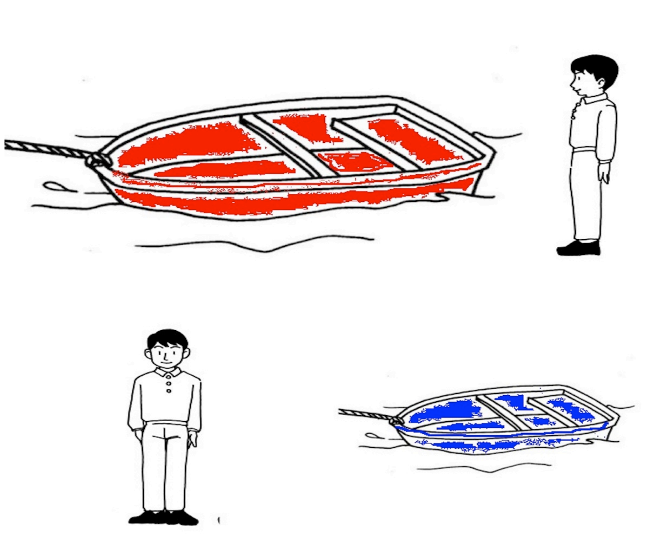
何が赤いですか？

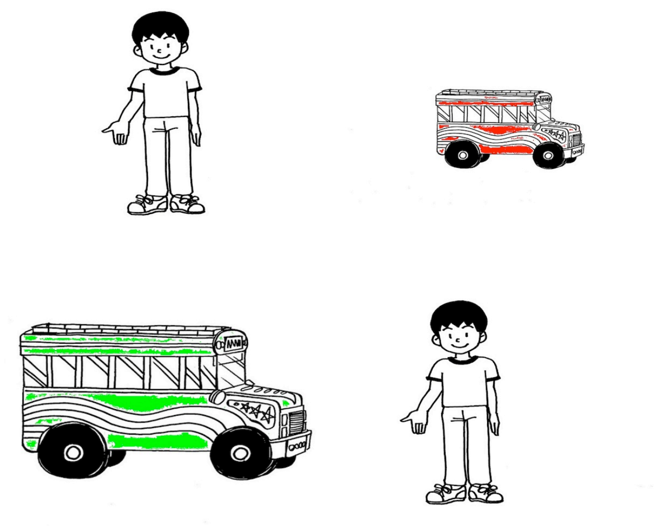
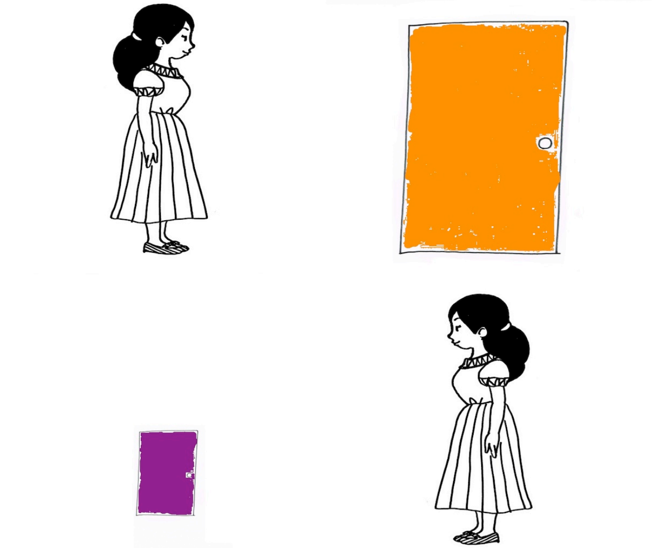
(Nani-ga akai desu ka?)

What-NOM red is-Q

→ “What is red?”



<u>1. Top right:</u> *a. Spanish:* ¿Qué es gris? What is grey *b. Japanese:* 何が灰色ですか？ (Nani-ga haiiro desu ka?) What-NOM grey is-Q → "What is grey?"	<u>2. Bottom left:</u> *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → "What is green?"	
<u>1. Top left:</u> *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → "What is red?"	<u>2. Bottom right:</u> *a. Spanish:* ¿Qué es azul? What is blue *b. Japanese:* 何が青いですか？ (Nani-ga aoi desu ka?) What-NOM blue is-Q → "What is blue?"	

<u>1. Top right:</u> *a. Spanish:* ¿Qué es rojo? What is red *b. Japanese:* 何が赤いですか？ (Nani-ga akai desu ka?) What-NOM red is-Q → “What is red?”	<u>2. Bottom left:</u> *a. Spanish:* ¿Qué es verde? What is green *b. Japanese:* 何が緑ですか？ (Nani-ga midori desu ka?) What-NOM green is-Q → “What is green?”	
<u>1. Top right:</u> *a. Spanish:* ¿Qué es naranja? What is orange *b. Japanese:* 何がオレンジですか？ (Nani-ga orenji desu ka?) What-NOM orange is-Q → “What is orange?”	<u>2. Top right:</u> *a. Spanish:* ¿Qué es morado? What is purple *b. Japanese:* 何が紫ですか？ (Nani-ga murasaki desu ka?) What-NOM purple is-Q → “What is purple?”	

Appendix 3: Statistical data Experiment 1 and Experiment 2

Experiment 1.
Spanish monolinguals. Analysis 1
<u>All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)</u>

	TW1–0-350		**TW2–400-1000**		**TW3–1000-1800**		**TW4–1800-2500**		**TW5–2500-4500**		**TW6–4500-6000**	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Patient HN	0.17211	0.15287	0.15436	0.59517	0.25745	0.50035	0.28974	0.47382	0.38479	0.35459	0.34299	0.34380
Agent HN	0.21201	0.10849	0.55257	0.17927	0.51854	0.21606	0.39094	0.32619	0.31134	0.39328	0.34685	0.32760
Estimate	0.03626	-0.0424	0.39126	-0.41032	0.25635	-0.2837	0.09624	-0.1374	-0.0546	0.01679	0.00818	-0.0236
Std. Error	0.03789	0.02757	0.04083	0.03849	0.04292	0.03907	0.04716	0.04277	0.04462	0.04231	0.03555	0.057
t value	0.957	-1.539	9.582	-10.66	5.972	-7.262	2.04	-3.212	-1.225	0.397	0.23	-0.414
p value	0.3465	0.1279	3.129 e-13 ***	1.557 e-13 ***	6.611 e-07 ***	3.82 e-09 ***	0.04676 *	0.00248 **	0.2254	0.6934	0.8193	0.6794
AA	0.20623	0.10499	0.36121	0.41079	0.38221	0.39646	0.34385	0.43754	0.34715	0.40971	0.37982	0.31569
AI	0.15173	0.10233	0.44141	0.28833	0.45869	0.26368	0.40882	0.32498	0.40467	0.29470	0.42990	0.24443
IA	0.22053	0.17684	0.26601	0.46149	0.32634	0.41809	0.27279	0.44265	0.29406	0.42304	0.24442	0.43111
Estimate	0.01525	0.03154	-0.0569	0.03234	-0.03485	0.02254	-0.0376	0.00691	-0.0301	0.00996	-0.0719	0.05903
Std. Error	0.02261	0.01653	0.0206	0.0206	0.01834	0.02017	0.02166	0.02182	0.01649	0.01803	0.01892	0.01884
t value	0.674	1.908	-2.763	1.57	-1.9	1.117	-1.738	0.317	-1.826	0.553	-3.803	3.133
p value	0.506	0.06346 △	0.00938 **	0.1248	0.06321 △	0.2696	0.08929 △	0.753	0.07167 △	0.5818	0.00038 ***	0.00298 **
Pat-HN AA	0.18988	0.11586	0.14818	0.64057	0.25298	0.52976	0.31047	0.49830	0.39828	0.37054	0.39089	0.31512
Pat-HN AI	0.14001	0.13288	0.22429	0.51823	0.32819	0.38923	0.36286	0.39903	0.42702	0.29225	0.42799	0.26216
Pat-HN IA	0.18787	0.20450	0.09167	0.63083	0.19249	0.58305	0.20075	0.52544	0.33236	0.40102	0.22999	0.43736
Ag-HN AA	0.22600	0.09184	0.61684	0.13504	0.53730	0.23651	0.38390	0.36463	0.28581	0.45671	0.36466	0.31649
Ag-HN AI	0.16254	0.07416	0.64425	0.07355	0.58059	0.14638	0.45175	0.25579	0.38395	0.29698	0.43187	0.22604
Ag-HN IA	0.24997	0.15192	0.42410	0.30792	0.44772	0.26850	0.33812	0.36757	0.25911	0.44313	0.26003	0.42435
Estimate	0.00348	-0.0101	-0.0654	0.10334	-0.0164	-0.0021	0.02855	-0.0096	0.00742	0.00342	0.0243	0.00483
Std. Error	0.0529	0.03235	0.03745	0.03662	0.03493	0.03752	0.04248	0.04116	0.03293	0.03619	0.03801	0.03802
t value	0.066	-0.313	-1.747	2.822	-0.471	-0.057	0.672	-0.233	0.225	0.094	0.639	0.127
p value	0.9479	0.7544	0.086△	0.005**	0.6384	0.9544	0.5023	0.8161	0.8219	0.9248	0.5234	0.8989

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Experiment 1
Spanish monolinguals. Analysis 2:
Patient as HN only – Voice (Active / Passive) x Animacy (AA / AI / IA)

	TW1–0-350		TW2–400-1000		TW3–1000-1800		TW4–1800-2500		TW5–2500-4500		TW6–4500-6000	
	agent	patient	agent	patient	agent	patient	agent	patient	agent	patient	agent	patient
Active	0.17247	0.13237	0.18384	0.57781	0.28875	0.45292	0.33573	0.41521	0.37337	0.33111	0.31624	0.32271
Passive	0.17190	0.16459	0.13761	0.60504	0.23966	0.52730	0.26362	0.50712	0.39123	0.36784	0.35587	0.35396
Estimate	-0.00253	0.02116	-0.04311	0.06222	-0.028	0.06371	-0.0669	0.09686	-0.0039	0.0278	0.07159	-0.0181
Std. Error	0.03052	0.02787	0.0242	0.03913	0.03423	0.03979	0.03854	0.04213	0.03326	0.03413	0.04548	0.04328
t value	-0.083	0.759	-1.781	1.59	-0.818	1.601	-1.737	2.299	-0.119	0.814	1.574	-0.418
p value	0.9338	0.4503	0.07724 △	0.1203	0.4282	0.1122	0.08742	0.0231 *	0.9064	0.4216	0.1219	0.6798
AA	0.18988	0.11586	0.14818	0.64058	0.25298	0.52976	0.31047	0.49830	0.39828	0.37055	0.39089	0.31512
AI	0.14001	0.13288	0.22429	0.51823	0.32820	0.38923	0.36286	0.39904	0.42702	0.29225	0.42799	0.26216
IA	0.18787	0.20450	0.09168	0.63083	0.19249	0.58305	0.20075	0.52544	0.33236	0.40102	0.22999	0.43736
Estimate	0.00052	0.03829	-0.03039	-0.0085	-0.0313	0.02852	-0.0523	0.01422	-0.0279	0.00624	-0.0818	0.05726
Std. Error	0.02918	0.02105	0.02308	0.02558	0.02415	0.02985	0.03145	0.03281	0.02311	0.02864	0.02276	0.02319
t value	0.018	1.819	-1.317	-0.331	-1.296	0.955	-1.663	0.434	-1.206	0.218	-3.595	2.469
p value	0.9858	0.08029 △	0.1938	0.7409	0.1999	0.3424	0.1051	0.6654	0.2385	0.828	0.00111 **	0.01905 *
AA Active	0.22358	0.08145	0.13537	0.67325	0.31926	0.46115	0.40669	0.40798	0.38150	0.36411	0.39559	0.31783
AA Passive	0.17549	0.13055	0.15351	0.62699	0.22542	0.55829	0.27047	0.53585	0.40525	0.37322	0.38916	0.31412
AI Active	0.15206	0.11794	0.24449	0.49241	0.36306	0.33278	0.38542	0.37773	0.40454	0.28364	0.39583	0.29617
AI Passive	0.12651	0.14961	0.20168	0.54714	0.28918	0.45243	0.33760	0.42289	0.45184	0.30175	0.45903	0.22935
IA Active	0.16165	0.21178	0.11134	0.65346	0.11239	0.68202	0.16855	0.49624	0.30477	0.39143	0.09571	0.37469
IA Passive	0.19709	0.20194	0.08476	0.62286	0.22068	0.54823	0.21208	0.53571	0.34207	0.40440	0.27080	0.45640
Estimate	0.04714	-0.0319	-0.0124	0.00586	0.10989	-0.12869	0.08329	-0.03753	-0.0019	-0.0013	0.09188	0.02868
Std. Error	0.03942	0.03115	0.02818	0.03963	0.0359	0.04034	0.04114	0.04611	0.03374	0.04049	0.04661	0.04096
t value	1.196	-1.025	-0.44	0.148	3.061	-3.19	2.024	-0.814	-0.059	-0.033	1.971	0.7
p value	0.235 0.235	0.3067	0.6604	0.8825	0.003**	0.0016**	0.0452*	0.4182	0.9536	0.9738	0.0494*	0.4846

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Experiment 2
Japanese monolinguals. Analysis 1:
<u>All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)</u>

	TW1–0-350		TW2–400-1000		TW3–1000-1800		TW4–1800-2500		TW5–2500-4500		TW6–4500-6000	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Patient HN	0.19499	0.28724	0.30657	0.59700	0.54316	0.37090	0.48979	0.39908	0.30833	0.57983	0.20721	0.62869
Agent HN	0.30349	0.17212	0.57368	0.36342	0.35679	0.56710	0.38577	0.50463	0.57674	0.30833	0.67571	0.18552
Estimate	0.13406	-0.1157	0.26187	-0.2295	-0.1703	0.19614	-0.1194	0.11467	0.26123	-0.26331	0.46543	-0.4386
Std. Error	0.03018	0.03203	0.03859	0.03366	0.0449	0.05065	0.04756	0.04735	0.05371	0.05313	0.04064	0.04015
t value	4.442	-3.611	6.787	-6.819	-3.792	3.872	-2.51	2.422	4.864	-4.956	11.45	-10.92
p value	6.905 e-05 ***	0.00059 ***	4.207 e-09 ***	5.053 e-09 ***	0.00069 ***	0.00035 ***	0.01602 *	0.0189 *	1.438 e-05 ***	1.261 e-05 ***	2.039 e-14 ***	3.795 e-13 ***
AA	0.22150	0.27314	0.43391	0.50735	0.44026	0.50804	0.43942	0.46959	0.44256	0.46971	0.46163	0.40054
AI	0.25644	0.17156	0.53763	0.36699	0.46095	0.42075	0.49665	0.39221	0.54510	0.33999	0.52072	0.33612
IA	0.27459	0.24166	0.36492	0.55608	0.43322	0.49573	0.36930	0.50204	0.35485	0.50725	0.36776	0.46054
Estimate	0.04029	-0.0222	-0.0119	0.01493	-0.0029	-0.0035	-0.0293	0.01637	-0.0346	0.00831	-0.0361	0.02045
Std. Error	0.02337	0.02157	0.02316	0.02748	0.02093	0.02045	0.02162	0.02049	0.01841	0.01796	0.01764	0.01631
t value	1.724	-1.035	-0.517	0.543	-0.14	-0.171	-1.356	0.799	-1.88	0.463	-2.048	1.254
p value	0.09582 △	0.3018	0.6378	0.5877	0.9038	0.8645	0.2089	0.4257	0.0648 △	0.6441	0.04494 *	0.2144
Pat-HN AA	0.14731	0.35702	0.25881	0.66313	0.58935	0.35826	0.51717	0.38748	0.30949	0.60724	0.20186	0.65315
Pat-HN AI	0.16642	0.22029	0.40917	0.48018	0.53685	0.33818	0.54612	0.34626	0.38262	0.49670	0.24908	0.58965
Pat-HN IA	0.26235	0.28854	0.25544	0.64607	0.50560	0.41410	0.41282	0.45816	0.23817	0.63153	0.17398	0.64208
Ag-HN AA	0.28613	0.20007	0.58467	0.37184	0.30923	0.64103	0.36821	0.54479	0.56032	0.34800	0.69543	0.17319
Ag-HN AI	0.33287	0.13019	0.64606	0.27040	0.39813	0.48938	0.45567	0.43028	0.68268	0.20729	0.75048	0.12169
Ag-HN IA	0.28776	0.19117	0.47925	0.46146	0.36050	0.57778	0.32406	0.54764	0.47489	0.37939	0.57072	0.27040
Estimate	-0.0402	0.0259	-0.0543	0.05706	0.06134	-0.0611	0.03729	-0.0426	-0.0127	0.00600	-0.0539	0.05397
Std. Error	0.04048	0.0389	0.03902	0.04043	0.03835	0.03895	0.04263	0.04002	0.03569	0.03489	0.0333	0.03089
t value	-0.992	0.666	-1.391	1.411	1.599	-1.568	0.875	-1.066	-0.356	0.172	-1.618	1.747
p value	0.3321	0.5062	0.1669	0.161	0.116	0.1204	0.5486	0.288	0.7221	0.8634	0.1089	0.08375

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Experiment 2
Japanese monolinguals. Analysis 2:
RCs with patient as HN only– Voice (Active / Passive) x Animacy (AA / AI / IA)

	TW1–0-350		TW2–400-1000		TW3–1000-1800		TW4–1800-2500		TW5–2500-4500		TW6–4500-6000	
	agent	patient	agent	patient	agent	patient	agent	patient	agent	patient	agent	patient
Active	0.15507	0.20985	0.36415	0.53819	0.52803	0.37164	0.53721	0.34893	0.43845	0.44507	0.33517	0.52021
Passive	0.20775	0.31197	0.28834	0.61606	0.54805	0.37065	0.47481	0.41492	0.26655	0.62311	0.16415	0.66519
Estimate	0.05164	0.07548	-0.02973	0.02572	-0.01683	-0.00410	-0.02076	0.03508	-0.12355	0.13555	-0.13397	0.12812
Std. Error	0.03690	0.03768	0.03608	0.03505	0.03970	0.03838	0.04520	0.04254	0.03309	0.03397	0.03340	0.03711
t value	1.399	2.003	-0.824	0.734	-0.424	-0.107	-0.459	0.825	-3.734	3.99	-4.011	3.453
p value	0.1621	0.04996 *	0.4221	0.478	0.6736	0.9157	0.6525	0.4258	0.00036 ***	0.00016 ***	9.84e-05 ***	0.00071 ***
AA	0.14731	0.35702	0.25881	0.66313	0.58935	0.35826	0.51717	0.38748	0.30949	0.60724	0.20186	0.65315
AI	0.16642	0.22029	0.40917	0.48018	0.53685	0.33818	0.54612	0.34626	0.38262	0.49670	0.24908	0.58965
IA	0.26235	0.28854	0.25544	0.64607	0.50560	0.41410	0.41282	0.45816	0.23817	0.63153	0.17398	0.64208
Estimate	0.06537	-0.0355	0.00476	-0.0136	-0.0429	0.02819	-0.0543	0.03788	-0.0289	0.00613	-0.0085	-0.0059
Std. Error	0.02709	0.02853	0.02808	0.02891	0.02742	0.0273	0.03044	0.02545	0.02339	0.02388	0.02149	0.02056
t value	2.413	-1.244	0.169	-0.47	-1.568	1.032	-1.785	1.489	-1.238	0.257	-0.394	-0.286
p value	0.01974 *	0.2177	0.8655	0.6388	0.1242	0.3064	0.08059 △	0.1427	0.223	0.7977	0.6942	0.7748
AA Active	0.06349	0.37075	0.23056	0.69345	0.49955	0.44719	0.43687	0.44999	0.45624	0.46001	0.50996	0.41642
AA Passive	0.15920	0.35507	0.26246	0.65909	0.60129	0.34601	0.52784	0.37917	0.29042	0.62637	0.15672	0.68783
AI Active	0.17996	0.18044	0.40036	0.49355	0.51730	0.36630	0.55844	0.32252	0.42612	0.44634	0.29174	0.54737
AI Passive	0.14830	0.27359	0.42158	0.46105	0.56484	0.29707	0.52873	0.37979	0.31925	0.57006	0.18300	0.65515
IA Active	0.07500	0.16607	0.21554	0.74187	0.76549	0.23050	0.52584	0.43133	0.56922	0.38343	0.42991	0.44509
IA Passive	0.27033	0.29376	0.25701	0.64218	0.49521	0.42171	0.40837	0.45922	0.22531	0.64116	0.16326	0.65033
Estimate	0.03263	0.09369	0.00489	-0.0145	-0.1135	0.09144	-0.11943	0.0872	-0.05251	0.025284	0.09091	-0.07075
Std. Error	0.06516	0.06742	0.05766	0.06044	0.06566	0.0633	0.07645	0.0752	0.056	0.05889	0.0573	0.06848
t value	0.501	1.39	0.085	-0.24	-1.728	1.445	-1.562	1.16	-0.938	0.429	1.587	-1.033
p value	0.6173	0.1657	0.9333	0.8128	0.08598	0.154	0.1194	0.2476	0.3501	0.6693	0.1132	0.302

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Appendix 4: Statistical data Experiment 3

Spanish-Japanese Bilingual speakers. Analysis 1: All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)
Analysis with all valid responses (correct and incorrect)

	TW1 – 0-400 ms.		**TW2 – 400-800 ms**		**TW3 – 800-1200 ms.**		**TW4 – 1200-1600 ms,**		**TW5 – 1600-2000 ms.**		**TW6 – 2000-2400 ms.**		**TW7 – 2400-2800 ms**		**TW8 – 2800-3200 ms**		**TW9 – 3200-3600 ms**	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Pat HN	0.11600	0.17505	0.15890	0.38437	0.21166	0.32101	0.24248	0.27108	0.31484	0.22448	0.38905	0.19477	0.36070	0.21474	0.40138	0.19083	0.33111	0.25
Agent HN	0.19708	0.08552	0.38187	0.18177	0.27177	0.28635	0.23865	0.30115	0.21594	0.34094	0.25489	0.33771	0.29365	0.33344	0.24792	0.38198	0.30437	0.35146
Estimate	0.08765	-0.0942	0.21692	-0.21121	0.06325	-0.04695	-0.00132	0.00739	-0.09892	0.10148	-0.14987	-0.14987	-0.08274	0.12344	-0.16488	0.18359	-0.05259	1.05E-01
Std. Error	0.03043	0.02656	0.05657	0.05235	0.0335	0.03817	0.03262	0.05016	0.0341	0.04868	0.06444	0.06444	0.04212	0.05642	0.04255	0.05722	0.06276	6.76E-02
t value	2.88	-3.547	3.834	-4.035	1.888	-1.23	-0.041	0.147	-2.901	2.085	-2.326	-2.326	-1.964	2.188	-3.875	3.208	-0.838	1.553
p value	0.00519 **	0.00074 ***	0.00209 **	0.00084 ***	0.09046	0.28	0.9006	0.8024	0.00285 **	0.05745 △	0.04352 *	0.04036 *	0.05608 △	0.04294 *	0.00028 ***	0.00927 **	0.4686	0.1434
AA	0.16783	0.108	0.28133	0.2785	0.22883	0.33333	0.21667	0.3325	0.19967	0.33233	0.32033	0.28	0.3065	0.32633	0.3005	0.33317	0.333	0.329
AI	0.14615	0.09415	0.33524	0.20340	0.22901	0.30391	0.25572	0.25604	0.32490	0.23887	0.34812	0.25254	0.36896	0.23680	0.36593	0.25731	0.37786	0.26415
IA	0.16722	0.18455	0.21368	0.35269	0.27894	0.26585	0.24889	0.27378	0.25553	0.29462	0.27507	0.28650	0.29240	0.27581	0.28245	0.29462	0.22751	0.32577
Estimate	0.00585	0.02913	-0.01938	0.02271	0.02853	-0.04044	0.01587	-0.03182	0.0242	-0.01325	-0.02855	-0.02855	-0.01303	-0.01522	-0.01771	-0.00913	-0.04798	9.46E-05
Std. Error	0.01879	0.01642	0.02093	0.02344	0.02076	0.02360	0.02025	0.02028	0.02115	0.0205	0.02174	0.02174	0.02606	0.02356	0.02633	0.02242	0.02379	2.58E-02
t value	0.311	1.775	-0.926	0.969	1.375	-1.713	0.784	-1.569	1.144	-0.647	-1.314	-1.314	-0.5	-0.646	-0.673	-0.408	-2.017	0.004
p value	0.9	0.08985 △	0.3864	0.3182	0.2245	0.1068	0.4176	0.1183	0.1628	0.4739	0.2158	0.5896	0.7039	0.4565	0.6612	0.6432	0.05249 △	0.9761
Pat-HN AA	0.13702	0.14543	0.10657	0.41466	0.16226	0.35537	0.18990	0.31931	0.24839	0.26362	0.44952	0.13982	0.37540	0.22716	0.39343	0.19431	0.34575	0.26202
Pat-HN AI	0.09905	0.13633	0.23355	0.28472	0.21199	0.32785	0.27814	0.21564	0.40570	0.15022	0.38633	0.24671	0.37134	0.23721	0.46820	0.13596	0.40607	0.16447
Pat-HN IA	0.11389	0.2375	0.13333	0.45278	0.25417	0.28472	0.25417	0.28194	0.28611	0.26111	0.33924	0.19305	0.33785	0.18264	0.34479	0.23993	0.24722	0.32083
Ag-HN AA	0.18978	0.08134	0.40582	0.18151	0.27626	0.31764	0.23573	0.34189	0.16495	0.38128	0.22831	0.37985	0.25742	0.39697	0.23430	0.43208	0.32391	0.37671
Ag-HN AI	0.18243	0.06165	0.41357	0.14077	0.24212	0.28547	0.23846	0.28716	0.26267	0.30715	0.31869	0.25704	0.36712	0.23649	0.28716	0.35079	0.35614	0.34093
Ag-HN IA	0.22759	0.12461	0.30464	0.23939	0.30699	0.24449	0.24292	0.26454	0.22091	0.33255	0.20244	0.39229	0.24096	0.38129	0.21187	0.35652	0.20519	0.33137
Estimate	0.03893	-0.0241	-0.05648	0.01042	-0.03256	0.00960	-0.02662	-0.02377	0.01698	-0.02772	0.04667	0.04667	0.01804	0.00908	0.01838	-0.06404	-0.00792	-4.73E-02
Std. Error	0.03752	0.03278	0.04187	0.04687	0.04149	0.04715	0.04053	0.04059	0.04229	0.041	0.04355	0.04355	0.05205	0.04712	0.05261	0.04485	0.04764	5.17E-02
t value	1.038	-0.735	-1.349	0.222	-0.785	0.204	-0.657	-0.586	0.402	-0.676	1.072	1.072	0.346	0.193	0.349	-1.428	-0.166	-0.915
p value	0.3009	0.4648	0.1807	0.8243	0.4344	0.8388	0.5119	0.5584	0.688	0.4995	0.2843	0.5361	0.7292	0.8472	0.7277	0.1605	0.8682	0.3631

	TW10 – 3600-4000 ms.		TW11 – 4000-4400 ms		TW12 – 4400-4800 ms.		TW13 – 4800-5200 ms,		TW14 – 5200-5600 ms.		TW15 – 5600-6000 ms.		TW16 – 6000-6400 ms		TW17 – 6400-6800 ms		TW18 – 6800-7200 ms	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Pat HN	0.30646	0.26516	0.30177	0.28772	0.25394	0.33247	0.25838	0.32199	0.23336	0.34960	0.22177	0.35714	0.20246	0.38617	0.19119	0.34825	0.16233	0.38868
Agent HN	0.32021	0.30885	0.32297	0.28141	0.33794	0.26047	0.38484	0.21744	0.40969	0.19854	0.37225	0.20942	0.35125	0.20177	0.38125	0.15626	0.35042	0.13758
Estimate	0.00139	0.05308	-0.02239	-0.00184	0.04358	-0.07983	0.11879	-0.11124	0.16044	-0.14831	0.11979	-0.13374	0.120174	-0.17853	0.17409	-0.18879	0.18971	-0.24144
Std. Error	0.07374	0.06116	0.06594	0.04518	0.07358	0.04427	0.04289	0.04183	0.05953	0.07869	0.06945	0.07325	0.06548	0.03708	0.05942	0.03465	0.03598	0.05569
t value	0.019	0.868	-0.34	-0.041	0.592	-1.803	2.769	-2.66	2.695	-1.885	1.725	-1.826	1.836	-4.815	2.93	-5.448	5.273	-4.335
p value	0.817	0.4115	0.7695	0.8888	0.576	0.08343 △	0.01186 *	0.01662 *	0.0235 *	0.08121 △	0.1131	0.08597 △	0.08369 △	1.21e-05 ***	0.01164 *	3.221e-0 ***	1.807e-0 ***	0.00153 **
AA	0.32567	0.2885	0.31633	0.27483	0.245	0.34983	0.26967	0.33283	0.30517	0.28133	0.30517	0.2605	0.30617	0.266	0.318	0.2355	0.27967	0.27
AI	0.34732	0.27560	0.38029	0.23061	0.38285	0.22660	0.39054	0.21362	0.37611	0.23537	0.32125	0.27742	0.31113	0.25709	0.33985	0.22362	0.30802	0.22837
IA	0.26217	0.30457	0.23267	0.35656	0.26346	0.30826	0.31656	0.25055	0.30052	0.29019	0.28053	0.29501	0.22511	0.34237	0.21497	0.27765	0.19653	0.26139
Estimate	-0.01820	0.00806	-0.03883	0.041957	0.0174	-0.0239	0.03373	-0.04801	0.01395	-0.00558	-0.00208	0.01273	-0.02982	0.03004	-0.03819	0.010663	-0.02722	-0.01539
Std. Error	0.028511	0.02282	0.028531	0.027921	0.02975	0.02737	0.0265	0.02588	0.02254	0.021542	0.02274	0.02204	0.023677	0.02297	0.02206	0.021462	0.02226	0.02216
t value	-0.639	0.353	-1.361	1.503	0.585	-0.873	1.273	-1.855	0.619	-0.259	-0.091	0.578	-1.26	1.308	-1.731	0.497	-1.223	-0.695
p value	0.5477	0.7356	0.178	0.1291	0.581	0.4809	0.3372	0.1315	0.6245	0.8579	0.8727	0.5098	0.1939	0.1832	0.07012	0.4357	0.1865	0.5747
Pat-HN AA	0.3125	0.22195	0.34295	0.24439	0.21034	0.37820	0.22957	0.37419	0.23718	0.34695	0.23518	0.30088	0.21554	0.38862	0.18349	0.34135	0.18790	0.41106
Pat-HN AI	0.30848	0.25656	0.32237	0.25073	0.33735	0.26645	0.31798	0.25658	0.27449	0.31506	0.21820	0.40387	0.24137	0.36169	0.23757	0.33260	0.18781	0.36001
Pat-HN IA	0.29931	0.31076	0.24653	0.36042	0.2125	0.35556	0.22674	0.33889	0.19097	0.38472	0.21354	0.36151	0.15417	0.40729	0.15382	0.36909	0.11597	0.39653
Ag-HN AA	0.33505	0.33590	0.29737	0.29652	0.26969	0.32962	0.29823	0.30337	0.35359	0.23459	0.35502	0.23173	0.37072	0.17865	0.41381	0.16010	0.34503	0.16952
Ag-HN AI	0.37725	0.29026	0.42551	0.21489	0.41838	0.19549	0.44720	0.18008	0.45439	0.17399	0.40062	0.18001	0.36486	0.17652	0.41864	0.13968	0.40062	0.12697
Ag-HN IA	0.22013	0.29756	0.21698	0.35220	0.32115	0.25472	0.41824	0.15055	0.42453	0.18318	0.35636	0.21973	0.30542	0.26887	0.28419	0.17413	0.28773	0.10841
Estimate	-0.04983	-0.06171	0.00421	-0.03294	0.02079	-0.0302	0.06504	-0.05754	0.06639	-0.04198	0.0149	-0.03749	0.00168	0.03179	-0.04431	-0.00961	0.01532	-0.02652
Std. Error	0.05363	0.04569	0.05704	0.05577	0.05948	0.05468	0.05294	0.05173	0.04513	0.04316	0.04554	0.04415	0.04739	0.04592	0.04418	0.04291	0.0445	0.04439
t value	-0.929	-1.351	0.074	-0.591	0.35	-0.552	1.229	-1.112	1.471	-0.973	0.327	-0.849	0.035	0.692	-1.003	-0.224	0.344	-0.597
p value	0.3636	0.1824	0.9413	0.5565	0.7275	0.5821	0.2263	0.2741	0.1432	0.3379	0.7437	0.3963	0.9718	0.4889	0.3164	0.822	0.7315	0.5509

Experiment 3.
Spanish-Japanese Bilingual speakers. Analysis 2 (TWs grouped)
<u>All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)</u>
<u>Analysis with all valid responses (correct and incorrect)</u>

	TW1 – 0-350 ms.		**TW2 – 400-1000 ms**		**TW3 – 1500-3200 ms.**		**TW4 – 3200-4800 ms.**		**TW5 – 4800-7200 ms**	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Pat HN	0.12003	0.14299	0.17069	0.36428	0.35793	0.20944	0.29832	0.28384	0.21282	0.36458
Agent HN	0.17476	0.07988	0.35339	0.20910	0.25067	0.34676	0.32266	0.29987	0.383565	0.19307
Estimate	0.06262	-0.07054	0.17418	-0.16145	-0.12397	0.12737	-0.01829	0.02686	0.14433	-0.16438
Std. Error	0.03207	0.02757	0.04659	0.03273	0.04473	0.04768	0.06623	0.05929	0.05627	0.05526
t value	1.953	-2.558	3.739	-4.932	-2.772	2.671	-0.276	0.453	2.565	-2.974
p value	0.05302 .	0.01215 *	0.00328 **	5.149e-06 ***	0.01499 *	0.02157 *	0.7826	0.6511	0.02391 *	0.01077 *
AA	0.15695	0.092	0.26995	0.29138	0.27448	0.32038	0.305	0.31054	0.29977	0.27619
AI	0.12614	0.08288	0.30270	0.22793	0.34642	0.24489	0.37386	0.24837	0.34730	0.25036
IA	0.16898	0.15739	0.23122	0.32834	0.27374	0.28866	0.24645	0.32379	0.26261	0.29174
Estimate	0.00772	0.02758	-0.00572	0.00959	-0.00258	-0.00783	-0.02189	0.00763	-0.00742	0.00119
Std. Error	0.02027	0.01726	0.01812	0.02371	0.01557	0.01522	0.023	0.02237	0.01819	0.01734
t value	0.381	1.598	-0.316	0.404	-0.166	-0.514	-0.952	0.341	-0.408	0.068
p value	0.7034	0.1192	0.7532	0.687	0.8683	0.6089	0.3431	0.7331	0.6837	0.9454
Pat-HN AA	0.14011	0.11722	0.12796	0.39842	0.35201	0.21429	0.30288	0.27664	0.21503	0.36290
Pat-HN AI	0.09691	0.11069	0.22177	0.29442	0.40092	0.18989	0.34357	0.23456	0.24964	0.34556
Pat-HN IA	0.12460	0.19603	0.15919	0.40107	0.32222	0.22381	0.25139	0.33689	0.17591	0.38411
Ag-HN AA	0.16895	0.07404	0.37109	0.21514	0.21924	0.39596	0.30651	0.33469	0.36012	0.21442
Ag-HN AI	0.14865	0.06145	0.36504	0.17671	0.30444	0.28726	0.39719	0.25901	0.42356	0.17603
Ag-HN IA	0.21923	0.11366	0.31277	0.24601	0.21887	0.36208	0.24086	0.30896	0.36076	0.18716
Estimate	0.03933	-0.01915	-0.0398	0.01659	0.01863	-0.02623	-0.016	-0.04102	0.02356	-0.02389
Std. Error	0.03928	0.03292	0.03497	0.04033	0.030965	0.02987	0.04608	0.04429	0.03580	0.03420
t value	1.001	-0.582	-1.138	0.411	0.602	-0.878	-0.347	-0.926	0.658	-0.699
p value	0.3179	0.5625	0.2583	0.6814	0.5505	0.3828	0.7287	0.3578	0.513	0.4888

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Experiment 3

Spanish-Japanese Bilingual speakers. Analysis 1: All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)

Analysis with correct responses only

	TW1 – 0-400 ms.		**TW2 – 400-800 ms**		**TW3 – 800-1200 ms.**		**TW4 – 1200-1600 ms,**		**TW5 – 1600-2000 ms.**		**TW6 – 2000-2400 ms.**		**TW7 – 2400-2800 ms**		**TW8 – 2800-3200 ms**		**TW9 – 3200-3600 ms**	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Pat HN	0.11470	0.17392	0.20506	0.36751	0.24274	0.30454	0.29658	0.23478	0.38811	0.18984	0.48010	0.16948	0.41901	0.22191	0.45178	0.21138	0.36212	0.28933
Agent HN	0.18845	0.09574	0.38947	0.19831	0.23839	0.32506	0.22360	0.32427	0.20293	0.39488	0.26288	0.375	0.31616	0.36005	0.28308	0.39472	0.33794	0.38995
Estimate	0.073	-0.0832	0.18612	-0.1665	-0.0043	-0.0029	-0.0749	0.08797	-0.1875	0.21819	-0.2254	0.18392	-0.1114	0.14322	-0.1712	0.18516	-0.0171	0.10684
Std. Error	0.03687	0.03662	0.04223	0.0711	0.04407	0.07059	0.04306	0.04552	0.046	0.04696	0.04841	0.07063	0.05388	0.05419	0.055	0.05237	0.05485	0.08352
t value	1.98	-2.273	4.407	-2.341	-0.099	-0.042	-1.74	1.933	-4.075	4.647	-4.656	2.604	-2.069	2.643	-3.113	3.536	-0.311	1.279
p value	0.05315 △	0.02595 *	6.439e-05 ***	0.03504 *	0.9213	0.9676	0.08469	0.05656 △	9.376e-05 ***	3.335e-05 ***	6.458e-06 ***	0.03491 *	0.04103 *	0.00936 **	0.00274 **	0.00069 ***	0.7572	0.2175
AA	0.18434	0.11364	0.30966	0.28093	0.20297	0.38037	0.20865	0.35227	0.15593	0.42708	0.34407	0.32197	0.37342	0.36648	0.32449	0.37184	0.39204	0.39425
AI	0.13725	0.08119	0.36254	0.18728	0.22036	0.29081	0.26718	0.26224	0.33806	0.25666	0.38488	0.25172	0.37908	0.24334	0.37715	0.27062	0.39605	0.28608
IA	0.16521	0.22186	0.23977	0.38560	0.31689	0.28728	0.28070	0.25767	0.31652	0.27266	0.30044	0.32529	0.30336	0.33553	0.33845	0.34613	0.21418	0.40460
Estimate	-0.0051	0.04348	-0.0327	0.02957	0.05582	-0.0552	0.03632	-0.0533	0.08358	-0.0787	-0.0217	0.02052	-0.0430	-0.0108	-0.0014	-0.01013	-0.0857	0.00459
Std. Error	0.02437	0.02323	0.0312	0.03315	0.02815	0.03278	0.02836	0.02958	0.03082	0.03386	0.0332	0.03112	0.03594	0.03646	0.03819	0.03727	0.03457	0.03657
t value	-0.211	1.872	-1.047	0.892	1.983	-1.683	1.281	-1.801	2.712	-2.323	-0.655	0.659	-1.197	-0.295	-0.036	-0.272	-2.48	0.126
p value	0.8338	0.07014 △	0.2995	0.3736	0.04974 *	0.09594 △	0.2036	0.07533 △	0.00929 **	0.02694 *	0.5151	0.5115	0.2345	0.7708	0.9716	0.7878	0.01421 *	0.9004
Pat-HN AA	0.13225	0.14039	0.18116	0.35688	0.17301	0.29710	0.21286	0.27717	0.23460	0.23913	0.53804	0.07790	0.52446	0.20833	0.45018	0.18116	0.45018	0.30616
Pat-HN AI	0.08559	0.11430	0.27984	0.26013	0.23311	0.30799	0.31982	0.20495	0.48761	0.14527	0.50788	0.21453	0.41104	0.24718	0.50338	0.15822	0.45552	0.17173
Pat-HN IA	0.13793	0.27658	0.12859	0.51293	0.31034	0.30603	0.33333	0.23922	0.38290	0.20761	0.39871	0.18463	0.34555	0.20043	0.38721	0.30316	0.17313	0.42601
Ag-HN AA	0.21221	0.09932	0.37839	0.24031	0.21899	0.42490	0.20639	0.39244	0.11386	0.52762	0.24031	0.45252	0.29264	0.45107	0.25727	0.47384	0.36095	0.44138
Ag-HN AI	0.16910	0.06076	0.41354	0.14236	0.2125	0.28021	0.23472	0.29757	0.24583	0.32535	0.30903	0.27465	0.35937	0.24097	0.29931	0.33993	0.35937	0.35660
Ag-HN IA	0.19345	0.16518	0.35491	0.25372	0.32366	0.26786	0.22619	0.27678	0.24777	0.34003	0.19866	0.47098	0.25967	0.47545	0.28795	0.39062	0.25670	0.38244
Estimate	0.00238	-0.0418	0.0351	-0.0677	-0.0229	-0.0742	-0.0452	-0.0348	0.00794	-0.0761	0.06371	-0.0489	0.08704	0.00078	0.05924	-0.11522	0.10535	-0.0950
Std. Error	0.04721	0.04486	0.05534	0.06538	0.05725	0.06537	0.05728	0.0595	0.06079	0.06055	0.06372	0.06283	0.07028	0.07126	0.07182	0.06767	0.06969	0.07306
t value	0.05	-0.931	0.634	-1.036	-0.401	-1.136	-0.79	-0.584	0.131	-1.256	1	-0.778	1.239	0.011	0.825	-1.703	1.512	-1.3
p value	0.9599	0.3548	0.527	0.3024	0.6884	0.2622	0.4308	0.5594	0.8961	0.2101	0.3184	0.4374	0.2167	0.9913	0.4106	0.09389	0.1322	0.1963

	TW10 – 3600-4000 ms.		TW11 – 4000-4400 ms		TW12 – 4400-4800 ms.		TW13 – 4800-5200 ms,		TW14 – 5200-5600 ms.		TW15 – 5600-6000 ms.		TW16 – 6000-6400 ms		TW17 – 6400-6800 ms		TW18 – 6800-7200 ms	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Pat HN	0.30992	0.32280	0.31390	0.32959	0.29354	0.38179	0.25866	0.38343	0.20412	0.45389	0.17369	0.46375	0.16301	0.47088	0.14817	0.38624	0.12403	0.45833
Agent HN	0.375	0.33683	0.38622	0.28077	0.38077	0.27356	0.43910	0.21891	0.46231	0.18416	0.42232	0.19695	0.41714	0.17525	0.40522	0.15587	0.33954	0.16345
Estimate	0.02991	0.01403	0.0687	0.04846	0.09164	0.11176	0.18483	-0.167	0.2781	0.19695	0.2585	0.26765	0.26756	0.29563	0.22496	0.24489	0.21171	0.29312
Std. Error	0.0909	0.05159	0.05905	0.05628	0.06139	0.06033	0.05261	0.05701	0.05286	0.1093	0.05156	0.04907	0.04961	0.04827	0.06839	0.04453	0.05169	0.04894
t value	0.329	0.272	1.164	-0.861	1.493	-1.852	3.513	-2.929	5.261	-1.802	5.013	-5.454	5.393	-6.125	3.289	-5.5	4.096	-5.989
p value	0.7709	0.7856	0.2529	0.3932	0.14	0.0681	0.00092 ***	0.00485 **	3.763e-06 ***	0.1227	5.483e-06 ***	5.541e-07 ***	5.388e-07 ***	2.352e-08 ***	0.01335 *	8.793e-07 ***	0.00018 ***	9.569e-08 ***
AA	0.39646	0.35922	0.35574	0.31944	0.29324	0.39710	0.29482	0.37721	0.31976	0.31408	0.30997	0.30934	0.33018	0.30587	0.33018	0.23390	0.24779	0.33081
AI	0.36211	0.28759	0.39909	0.23589	0.40690	0.23568	0.41688	0.21354	0.39304	0.24141	0.33161	0.26813	0.33514	0.24132	0.35545	0.22018	0.30882	0.23067
IA	0.27047	0.37281	0.28691	0.38779	0.30190	0.36330	0.36184	0.30153	0.34210	0.35746	0.31856	0.36226	0.26060	0.37317	0.17544	0.31579	0.16155	0.31572
Estimate	-0.0528	0.00426	0.03736	0.03197	4.00E05	0.01792	0.03269	0.04669	0.02183	0.00987	0.00885	0.02205	0.03415	0.02592	0.05312	0.03848	0.03887	0.01415
Std. Error	0.03351	0.03392	0.03838	0.03679	4.00E02	0.03992	0.03747	0.03914	0.04174	0.03365	0.03923	0.03822	0.03683	0.03882	0.03222	0.03416	0.03738	0.03967
t value	-1.576	0.126	-0.973	0.869	0.001	-0.449	0.873	-1.193	0.523	-0.293	0.226	0.577	-0.927	0.668	-1.649	1.127	-1.04	-0.357
p value	0.1193	0.9001	0.3311	0.3855	0.9992	0.6548	0.3842	0.2342	0.6018	0.7695	0.8218	0.5653	0.3599	0.5072	0.1042	0.2696	0.3009	0.7223
Pat-HN AA	0.34783	0.35598	0.27627	0.39312	0.24456	0.47283	0.1875	0.49094	0.14855	0.53985	0.12591	0.41938	0.09420	0.54529	0.09692	0.37681	0.11685	0.53442
Pat-HN AI	0.33446	0.26520	0.35022	0.27027	0.38513	0.28998	0.33277	0.28322	0.26802	0.34065	0.20608	0.44200	0.23953	0.37083	0.21622	0.35529	0.15532	0.39471
Pat-HN IA	0.24856	0.36997	0.29741	0.35488	0.21552	0.42672	0.22055	0.42601	0.16667	0.53017	0.17026	0.52668	0.11997	0.53951	0.10201	0.43319	0.08979	0.47917
Ag-HN AA	0.42248	0.36095	0.39826	0.28004	0.31928	0.35659	0.35223	0.31638	0.41134	0.19331	0.40843	0.25048	0.45639	0.17781	0.45494	0.15746	0.31783	0.22190
Ag-HN AI	0.37917	0.30139	0.42973	0.21434	0.42055	0.20162	0.46963	0.16984	0.47014	0.18021	0.40903	0.16090	0.39409	0.16146	0.44132	0.13685	0.40347	0.12951
Ag-HN IA	0.29315	0.37574	0.27604	0.42187	0.39137	0.29762	0.50818	0.17262	0.52381	0.17857	0.47215	0.19196	0.40625	0.20089	0.25149	0.19419	0.23586	0.14643
Estimate	-0.0241	-0.0110	-0.0473	0.07143	0.07295	-0.0226	0.08201	-0.0466	0.06453	0.01037	0.01914	-0.0904	-0.0226	0.00337	0.08922	0.01854	0.01447	0.02064
Std. Error	0.06792	0.06906	0.07605	0.07357	0.07875	0.07808	0.06773	0.07248	0.06748	0.06727	0.06713	0.06528	0.06537	0.06460	0.06285	0.05825	0.06736	0.06485
t value	-0.355	-0.159	-0.622	0.971	0.926	-0.29	1.211	-0.642	0.956	0.154	0.285	-1.385	-0.347	0.052	-1.42	-0.318	-0.215	-0.318
p value	0.7246	0.8738	0.5364	0.3332	0.3565	0.7723	0.2273	0.5242	0.3433	0.8786	0.7763	0.167	0.7294	0.9584	0.1597	0.7503	0.8303	0.7503

Experiment 3
Spanish-Japanese Bilingual speakers. Analysis 2 (TWs grouped)
All RCs –RC type (Agent-HN / Patient-HN) x Animacy (AA / AI / IA)
Analysis with only correct responses

	TW1 – 0-350 ms.		TW2 – 400-1000 ms		TW3 – 1500-3200 ms.		TW4 – 3200-4800 ms.		TW5 – 4800-7200 ms	
	agent	patient	agent	patient	agent	patient	agent	patient	agent	patient
Pat HN	0.116104	0.148475	0.207433	0.347306	0.424023	0.198555	0.319874	0.330875	0.180958	0.447367
Agent HN	0.164486	0.089240	0.347132	0.233412	0.260414	0.378480	0.371750	0.319219	0.427694	0.191952
Estimate	0.04761	-0.06456	0.14024	-0.11481	-0.1723	0.19332	0.000356	0.05523	0.17703	-0.1942
Std. Error	0.03885	0.03603	0.03401	0.04451	0.03228	0.03559	0.082843	0.07354	0.06824	0.06786
t value	1.225	-1.792	4.124	-2.58	-5.338	5.431	0.004	0.751	2.594	-2.862
p value	0.2253	0.076 △	0.000134 ***	0.01183 *	4.593 e-07***	1.123 e-06***	0.9968	0.454	0.04842 *	0.06077 △
AA	0.171717	0.098124	0.275641	0.311577	0.284993	0.374675	0.359375	0.367503	0.310112	0.315333
AI	0.116102	0.072656	0.317209	0.210018	0.364163	0.253854	0.393324	0.260094	0.365274	0.250884
IA	0.162907	0.199666	0.262708	0.360549	0.310860	0.314035	0.268366	0.382127	0.28371500	0.348640
Estimate	-0.0002832	0.04247	-0.005015	0.009527	0.007432	-0.02103	-0.0608	0.005471	0.01025	-0.008666
Std. Error	0.025144	0.02252	0.024871	0.02955	0.02396	0.0275	0.03074	0.029641	0.02495	0.024113
t value	-0.011	1.886	-0.202	0.322	0.31	-0.765	-1.978	0.185	0.411	-0.359
p value	0.9911	0.06868 △	0.8412	0.7474	0.7629	0.456	0.04998 *	0.8539	0.6819	0.7199
Pat-HN AA	0.132505	0.119048	0.177257	0.342809	0.412422	0.188613	0.329710	0.382020	0.128830	0.489862
Pat-HN AI	0.081081	0.095882	0.256410	0.268191	0.467825	0.186615	0.381334	0.249296	0.241566	0.375626
Pat-HN IA	0.147783	0.238916	0.168877	0.451812	0.377340	0.221675	0.233657	0.394396	0.144972	0.505194
Ag-HN AA	0.192691	0.086932	0.328265	0.294872	0.216833	0.474197	0.375242	0.359738	0.407077	0.221980
Ag-HN AI	0.137698	0.058333	0.354701	0.174145	0.300239	0.295317	0.400716	0.266753	0.442854	0.172657
Ag-HN IA	0.178571	0.159014	0.359890	0.266026	0.242007	0.409694	0.304315	0.369420	0.427413	0.186494
Estimate	-0.003847	-0.03049	0.03192	-0.06181	0.04472	-0.05165	0.01262	-0.01825	0.015122	-0.03123
Std. Error	0.049744	0.0445	0.04489	0.05589	0.04249	0.04502	0.06211	0.06019	0.049291	0.04747
t value	-0.077	-0.685	0.711	-1.106	1.053	-1.147	0.203	-0.303	0.307	-0.658
p value	0.9386	0.4949	0.478	0.2721	0.2934	0.2578	0.8396	0.7621	0.76	0.5167

*** Below 0.001
** Below 0.01
* Below 0.05
△ Below 0.1

Experiment 3

Spanish-Japanese Bilingual speakers. Analysis 2: RCs with Patient as HN –Voice (Active / Passive) x Animacy (AA / AI / IA)

Analysis with all responses (correct and incorrect responses)

	TW1 – 0-400 ms.		TW2 – 400-800 ms		TW3 – 800-1200 ms.		TW4 – 1200-1600 ms,		TW5 – 1600-2000 ms.		TW6 – 2000-2400 ms.		TW7 – 2400-2800 ms		TW8 – 2800-3200 ms		TW9 – 3200-3600 ms	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Active	0.09947	0.16789	0.14379	0.37439	0.20282	0.31638	0.18811	0.30188	0.27308	0.24060	0.37765	0.19281	0.36622	0.19056	0.40278	0.15155	0.33619	0.19404
Passive	0.14117	0.18594	0.18190	0.39956	0.22512	0.32805	0.32525	0.22419	0.37842	0.19994	0.40640	0.19776	0.35230	0.25155	0.39925	0.25062	0.32338	0.33520
Estimate	0.0666	0.01357	-0.0313	0.09311	0.00819	0.00224	0.15171	-0.0870	0.08877	-0.0356	-0.1764	0.04009	-0.0643	0.05259	-0.0330	0.0936	-0.0383	0.10875
Std. Error	0.03525	0.03978	0.06963	0.06642	0.05664	0.05913	0.05258	0.06097	0.06132	0.05538	0.0716	0.0535	0.08155	0.05434	0.07034	0.04907	0.0654	0.0608
t value	1.889	0.341	-0.449	1.402	0.145	0.038	2.885	-1.427	1.448	-0.643	-2.463	0.749	-0.788	0.968	-0.47	1.907	-0.585	1.788
p value	0.06796 △	0.734	0.6662	0.1689	0.8855	0.971	0.00636 **	0.1551	0.1765	0.5236	0.02477 *	0.4693	0.4375	0.3394	0.6424	0.08474 △	0.5605	0.09603 △
AA	0.13702	0.14543	0.10657	0.41466	0.16226	0.35537	0.18990	0.31931	0.24840	0.26362	0.44952	0.13982	0.37540	0.22716	0.39343	0.19431	0.34575	0.26202
AI	0.09905	0.13633	0.23355	0.28472	0.21199	0.32785	0.27814	0.21564	0.40570	0.15022	0.38633	0.24671	0.37134	0.23721	0.46820	0.13596	0.40607	0.16447
IA	0.11389	0.2375	0.13333	0.45278	0.25417	0.28472	0.25417	0.28194	0.28611	0.26111	0.33924	0.19306	0.33785	0.18264	0.34479	0.23993	0.24722	0.32083
Estimate	-0.0125	0.0388	0.01366	0.01558	0.04695	-0.0380	0.02775	-0.0185	0.01538	0.00326	-0.0539	0.02609	-0.0128	-0.0228	-0.0219	0.02489	-0.0594	0.0239
Std. Error	0.02437	0.0257	0.02693	0.03196	0.03065	0.03228	0.03096	0.02942	0.03179	0.02799	0.03172	0.02637	0.0322	0.02776	0.03984	0.02892	0.04357	0.03751
t value	-0.512	1.51	0.507	0.487	1.532	-1.178	0.896	-0.627	0.484	0.116	-1.699	0.989	-0.398	-0.823	-0.55	0.861	-1.364	0.637
p value	0.6093	0.147	0.6131	0.6264	0.1315	0.2439	0.372	0.5308	0.6286	0.9074	0.09074	0.3235	0.6915	0.411	0.585	0.3966	0.1764	0.5277
AA Active	0.13086	0.17383	0.08008	0.4375	0.15430	0.37174	0.16016	0.35417	0.25325	0.29492	0.40365	0.1875	0.30599	0.24479	0.34570	0.19727	0.30338	0.21615
AA Passive	0.14687	0.1	0.14896	0.37812	0.175	0.32917	0.2375	0.26354	0.24062	0.21354	0.52292	0.06354	0.48646	0.19896	0.46979	0.18958	0.41354	0.33542
AI Active	0.08391	0.13542	0.19155	0.29630	0.20428	0.31944	0.21065	0.21296	0.33044	0.12269	0.39410	0.20891	0.40220	0.17130	0.48032	0.08970	0.35185	0.17766
AI Passive	0.125	0.13790	0.30556	0.26488	0.22520	0.34226	0.39385	0.22024	0.53472	0.19742	0.37302	0.31151	0.31845	0.35020	0.44742	0.21528	0.49901	0.14186
IA Active	0.08640	0.19669	0.15319	0.39767	0.24694	0.26103	0.19056	0.34681	0.23100	0.31434	0.33578	0.18076	0.38480	0.15993	0.37439	0.17402	0.35049	0.19056
IA Passive	0.14984	0.29086	0.10737	0.52484	0.26362	0.31570	0.33734	0.19711	0.35817	0.19151	0.34375	0.20913	0.27644	0.21234	0.30609	0.32612	0.11218	0.49119
Estimate	0.00649	0.07762	-0.0639	0.08238	-0.0085	0.04430	0.02386	-0.0376	0.05613	-0.0268	-0.0802	0.07331	-0.1558	0.04322	-0.1058	0.08055	-0.1984	0.10339
Std. Error	0.03799	0.03875	0.04388	0.06068	0.05311	0.06544	0.05594	0.05923	0.0644	0.05669	0.06215	0.05294	0.06392	0.05612	0.06905	0.05676	0.07121	0.05925
t value	0.171	2.003	-1.458	1.358	-0.159	0.677	0.427	-0.635	0.872	-0.473	-1.29	1.385	-2.437	0.77	-1.532	1.419	-2.786	1.745
p value	0.866	0.0469*	0.1485	0.176	0.874	0.4992	0.6707	0.5257	0.3843	0.6363	0.1993	0.1679	0.0167*	0.4417 0.4417	0.1277	0.1572	0.006**	0.083 △

	TW10 – 3600-4000 ms.		**TW11 – 4000-4400 ms**		**TW12 – 4400-4800 ms.**		**TW13 – 4800-5200 ms,**		**TW14 – 5200-5600 ms.**		**TW15 – 5600-6000 ms.**		**TW16 – 6000-6400 ms**		**TW17 – 6400-6800 ms**		**TW18 – 6800-7200 ms**	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Active	0.33231	0.19363	0.32067	0.24285	0.25184	0.27941	0.27941	0.26246	0.26961	0.25020	0.28595	0.25572	0.25928	0.29608	0.23468	0.30801	0.19485	0.32251
Passive	0.26710	0.37407	0.27301	0.35603	0.25715	0.41325	0.22637	0.41262	0.17817	0.50093	0.12407	0.51155	0.11598	0.52332	0.125	0.40951	0.11282	0.48941
Estimate	-0.0693	0.15891	-0.0440	0.08702	-0.0066	0.1012	-0.0701	0.14345	-0.0927	0.20844	-0.1666	0.23496	-0.1467	0.22724	-0.1062	0.09458	-0.0398	0.14665
Std. Error	0.06774	0.06172	0.07123	0.06041	0.06728	0.06787	0.06514	0.05901	0.05872	0.06883	0.05984	0.06711	0.05725	0.05876	0.05647	0.06821	0.04873	0.07359
t value	-1.023	2.575	-0.618	1.44	-0.098	1.491	-1.076	2.431	-1.578	3.028	-2.784	3.501	-2.563	3.867	-1.881	1.387	-0.818	1.993
p value	0.311	0.02247 *	0.5401	0.2237	0.9216	0.1741	0.2833	0.04377 *	0.1227	0.00643 **	0.00716 **	0.00150 **	0.01261 *	0.00032 ***	0.06789 △	0.1709	0.4441	0.05504 △
AA	0.3125	0.22195	0.34295	0.24439	0.21034	0.37820	0.22957	0.37420	0.23718	0.34695	0.23518	0.30088	0.21554	0.38862	0.18349	0.34134	0.18790	0.41106
AI	0.30848	0.25658	0.32237	0.25073	0.33735	0.26645	0.31798	0.25658	0.27449	0.31506	0.21820	0.40387	0.24137	0.36170	0.23757	0.33260	0.18781	0.36001
IA	0.29931	0.31077	0.24653	0.36042	0.2125	0.35556	0.22674	0.33889	0.19097	0.38472	0.21354	0.36151	0.15417	0.40729	0.15382	0.36910	0.11597	0.39653
Estimate	0.00769	0.04139	-0.0417	0.05872	0.00620	-0.0060	-0.0021	-0.0164	-0.0212	0.01713	-0.0101	0.03269	-0.0313	0.0106	-0.0141	0.01578	-0.0404	-0.0020
Std. Error	0.03392	0.03176	0.03702	0.04205	0.03845	0.03954	0.03636	0.03785	0.03012	0.03335	0.03174	0.03422	0.03092	0.03743	0.02887	0.03425	0.03872	0.03753
t value	0.227	1.303	-1.127	1.396	0.161	-0.151	-0.058	-0.433	-0.705	0.514	-0.319	0.955	-1.011	0.283	-0.487	0.461	-1.042	-0.054
p value	0.8211	0.1936	0.2678	0.1717	0.8722	0.8797	0.9536	0.6656	0.4813	0.6076	0.7501	0.3405	0.3127	0.7788	0.6266	0.6452	0.3091	0.9567
AA Active	0.34570	0.10872	0.41016	0.14388	0.20703	0.31575	0.26953	0.29167	0.29883	0.19466	0.33398	0.20703	0.34179	0.22396	0.25456	0.28059	0.26562	0.32682
AA Passive	0.25937	0.40312	0.23542	0.40521	0.21562	0.47812	0.16562	0.50625	0.13855	0.59062	0.07708	0.45104	0.01354	0.65208	0.06979	0.43854	0.06354	0.54583
AI Active	0.28241	0.23380	0.30961	0.24653	0.28241	0.26042	0.32697	0.23843	0.27893	0.28472	0.23900	0.34722	0.24849	0.34294	0.28009	0.29456	0.21643	0.27662
AI Passive	0.35317	0.29563	0.34425	0.25794	0.43155	0.27679	0.30258	0.28770	0.26686	0.36706	0.18254	0.50099	0.22917	0.39385	0.16468	0.39782	0.13875	0.50298
IA Active	0.37255	0.23100	0.24817	0.33211	0.26164	0.26532	0.23836	0.26042	0.23223	0.26593	0.29044	0.20466	0.19301	0.31434	0.16789	0.34804	0.10539	0.36703
IA Passive	0.20353	0.41506	0.24439	0.39744	0.14824	0.47356	0.21154	0.44151	0.13702	0.54006	0.11298	0.56662	0.10336	0.52885	0.13542	0.39663	0.12981	0.43510
Estimate	-0.0488	-0.0560	0.0644	-0.0719	-0.0918	0.03238	0.01947	-0.0097	0.02718	-0.0611	0.03443	0.05405	0.11037	-0.0979	0.07731	-0.023	0.07147	-0.0835
Std. Error	0.06214	0.06406	0.06607	0.06404	0.06009	0.06936	0.06257	0.06933	0.06084	0.06692	0.06337	0.06831	0.06136	0.07163	0.0579	0.001	0.05972	0.07534
t value	-0.786	-0.875	0.975	-1.124	-1.528	0.467	0.311	-0.141	0.447	-0.918	0.543	0.791	1.799	-1.367	1.335	-0.018	1.197	-1.108
p value	0.4323	0.3835	0.3406	0.2668	0.1292	0.6413	0.7563	0.8881	0.6552	0.3599	0.5871	0.4295 0.4295	0.073 △	0.173	0.1833	0.3565	0.2551	0.2695

Experiment 3

Spanish-Japanese Bilingual speakers. Analysis 2: RCs with Patient as HN –Voice (Active / Passive) x Animacy (AA / AI / IA)

Analysis with correct responses only

	TW1 – 0-400 ms.		TW2 – 400-800 ms		TW3 – 800-1200 ms.		TW4 – 1200-1600 ms,		TW5 – 1600-2000 ms.		TW6 – 2000-2400 ms.		TW7 – 2400-2800 ms		TW8 – 2800-3200 ms		TW9 – 3200-3600 ms	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Active	0.11574	0.13349	0.25617	0.30015	0.28318	0.25154	0.22377	0.29167	0.39583	0.19522	0.60494	0.11111	0.56481	0.13503	0.56559	0.09954	0.47068	0.15586
Passive	0.11425	0.19153	0.18280	0.39684	0.22513	0.32762	0.32829	0.21001	0.38474	0.18750	0.42574	0.19489	0.35551	0.25974	0.40222	0.26008	0.31485	0.34745
Estimate	-0.0002	0.05123	-0.0743	0.22445	-0.0399	0.06173	0.12478	-0.0702	0.01048	-0.0058	-0.2015	0.08378	-0.2093	0.12471	-0.1634	0.16054	-0.0899	0.17373
Std. Error	0.04789	0.05762	0.06330	0.08963	0.07847	0.08531	0.07809	0.07550	0.09006	0.06816	0.09441	0.06509	0.08643	0.07201	0.08548	0.07039	0.09992	0.08802
t value	-0.005	0.889	-1.173	2.504	-0.509	0.724	1.598	-0.929	0.116	-0.085	-2.134	1.287	-2.422	1.732	-1.911	2.281	-0.899	1.974
p value	0.9962	0.3928	0.2679	0.01951 *	0.6186	0.4788	0.1155	0.3928	0.9155	0.9431	0.03628 *	0.2001	0.02355 *	0.0859	0.06232	0.02459 *	0.4374	0.06454 △
AA	0.13225	0.14040	0.18116	0.35688	0.17301	0.29710	0.21286	0.27717	0.23460	0.23913	0.53804	0.07790	0.52446	0.20833	0.45018	0.18116	0.45018	0.30616
AI	0.08559	0.11430	0.27984	0.26014	0.23311	0.30800	0.31982	0.20495	0.48761	0.14527	0.50788	0.21453	0.41104	0.24718	0.50338	0.15822	0.45552	0.17173
IA	0.13793	0.27658	0.12859	0.51293	0.31034	0.30603	0.33333	0.23922	0.38290	0.20761	0.39871	0.18463	0.34555	0.20043	0.38721	0.30316	0.17313	0.42601
Estimate	0.00482	0.05675	-0.0330	0.06659	0.06453	-0.0068	0.05316	-0.0216	0.06684	-0.0127	-0.0735	0.04934	-0.0882	-0.0060	-0.0345	0.06506	-0.1377	0.06007
Std. Error	0.03516	0.03788	0.03938	0.04904	0.04180	0.04618	0.04345	0.04129	0.05282	0.04080	0.04912	0.03931	0.05244	0.04420	0.05232	0.04318	0.05076	0.04843
t value	0.137	1.498	-0.838	1.358	1.544	-0.148	1.223	-0.523	1.265	-0.312	-1.497	1.255	-1.682	-0.136	-0.659	1.507	-2.713	1.24
p value	0.891	0.1529	0.4033	0.1806	0.1274	0.8835	0.2333	0.6016	0.2079	0.7573	0.1372	0.2115	0.09508 △	0.8917	0.5118	0.1346	0.00784 **	0.2205
AA Active	0.16667	0.30729	0.29688	0.29167	0.21875	0.18229	0.03646	0.46354	0.14583	0.53646	0.57292	0.13021	0.73437	0.20312	0.48958	0.09375	0.69271	0.08333
AA Passive	0.12500	0.10526	0.15680	0.37061	0.16338	0.32127	0.25000	0.23794	0.25329	0.17654	0.53070	0.06689	0.48026	0.20943	0.44189	0.19956	0.39912	0.35307
AI Active	0.09320	0.08333	0.25219	0.27961	0.24342	0.28399	0.26316	0.18421	0.42654	0.10307	0.60088	0.12610	0.49890	0.13377	0.52851	0.09759	0.41228	0.17763
AI Passive	0.07755	0.14699	0.30903	0.23958	0.22222	0.33333	0.37963	0.22685	0.55208	0.18981	0.40972	0.30787	0.31829	0.36690	0.47685	0.22222	0.50116	0.16551
IA Active	0.17188	0.19792	0.23438	0.40625	0.53646	0.16667	0.22396	0.63021	0.50000	0.29167	0.65625	0.02083	0.70833	0.07292	0.81771	0.11458	0.52604	0.12500
IA Passive	0.13250	0.28917	0.11167	0.53000	0.27417	0.32833	0.35083	0.17667	0.36417	0.19417	0.35750	0.21083	0.28750	0.22083	0.31833	0.33333	0.11667	0.47417
Estimate	-0.0349	0.11760	0.00272	-0.0203	-0.1321	0.00008	-0.0533	-0.1245	-0.1443	0.13076	-0.1407	0.11988	-0.0807	0.06472	-0.2297	0.05824	-0.0815	0.04479
Std. Error	0.07234	0.07781	0.10276	0.11202	0.10956	0.12210	0.11731	0.11225	0.13784	0.11199	0.12932	0.10699	0.14199	0.11993	0.14026	0.11598	0.13918	0.13143
t value	-0.483	1.511	0.026	-0.181	-1.206	0.001	-0.455	-1.109	-1.047	1.168	-1.088	1.12	-0.569	0.54	-1.638	0.502	-0.586	0.341
p value	0.635	0.1372	0.9789	0.8598	0.2615	0.9995	0.6506	0.273	0.3057	0.2459	0.279	0.2642	0.57	0.5898	0.104	0.6158	0.5611	0.7334

	TW10 – 3600-4000 ms.		TW11 – 4000-4400 ms		TW12 – 4400-4800 ms.		TW13 – 4800-5200 ms,		TW14 – 5200-5600 ms.		TW15 – 5600-6000 ms.		TW16 – 6000-6400 ms		TW17 – 6400-6800 ms		TW18 – 6800-7200 ms	
	agent	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**	**agent**	**patient**
Active	0.39583	0.20293	0.45062	0.25926	0.39043	0.30170	0.33873	0.27238	0.30556	0.26312	0.34105	0.32870	0.31590	0.33380	0.24228	0.33179	0.21219	0.37500
Passive	0.27251	0.37500	0.25437	0.36022	0.25134	0.41667	0.22379	0.43179	0.15995	0.53696	0.10081	0.52256	0.09644	0.53058	0.10719	0.40995	0.08564	0.49462
Estimate	-0.0156	0.15755	-0.1733	0.10096	-0.0272	0.10890	-0.1112	0.16118	-0.1507	0.30477	-0.2559	0.19386	-0.2195	0.19678	-0.1361	0.07816	-0.1265	0.11962
Std. Error	0.14070	0.08975	0.09338	0.08735	0.11880	0.09236	0.08990	0.08732	0.07654	0.09757	0.07128	0.08774	0.06821	0.08884	0.06340	0.08119	0.05805	0.09166
t value	-0.111	1.755	-1.856	1.156	-0.229	1.179	-1.237	1.846	-1.969	3.124	-3.59	2.209	-3.217	2.215	-2.147	0.963	-2.18	1.305
p value	0.9153	0.1098	0.1075	0.2845	0.8268	0.244	0.2233	0.06751 △	0.06025	0.00263 **	0.00055 ***	0.02926 *	0.00283 **	0.02886 *	0.03432 *	0.337	0.03217 *	0.194
AA	0.34783	0.35598	0.27627	0.39312	0.24457	0.47283	0.18750	0.49094	0.14855	0.53986	0.12591	0.41938	0.09420	0.54529	0.09692	0.37681	0.11685	0.53442
AI	0.33446	0.26520	0.35023	0.27027	0.38514	0.28998	0.33277	0.28322	0.26802	0.34065	0.20608	0.44200	0.23953	0.37083	0.21622	0.35529	0.15532	0.39471
IA	0.24856	0.36997	0.29741	0.35489	0.21552	0.42672	0.22055	0.42601	0.16667	0.53017	0.17026	0.52668	0.11997	0.53951	0.10201	0.43319	0.08980	0.47917
Estimate	-0.0343	-0.0007	0.01203	-0.0147	-0.0269	-0.0139	0.00271	-0.0287	0.00793	-0.0211	0.01776	0.05515	0.00425	0.00541	-0.0031	0.03059	-0.0167	-0.0222
Std. Error	0.04754	0.05042	0.05190	0.05366	0.05548	0.05933	0.04885	0.05715	0.04489	0.05881	0.04211	0.05410	0.04274	0.05510	0.03904	0.04917	0.03581	0.05582
t value	-0.722	-0.013	0.232	-0.274	-0.485	-0.236	0.055	-0.502	0.177	-0.358	0.422	1.019	0.099	0.098	-0.079	0.622	-0.466	-0.398
p value	0.4746	0.9894	0.8171	0.7856	0.6299	0.814	0.9559	0.6166	0.8606	0.7221	0.6734	0.3094	0.9208	0.9219	0.9367	0.5343	0.6421	0.6909
AA Active	0.75521	0.03125	0.59896	0.23438	0.32813	0.46354	0.25000	0.29167	0.16146	0.15104	0.33854	0.23438	0.47396	0.08333	0.20833	0.20833	0.35417	0.59375
AA Passive	0.26206	0.42434	0.20833	0.42654	0.22697	0.47478	0.17434	0.53289	0.14583	0.62171	0.08114	0.45833	0.01425	0.64254	0.07346	0.41228	0.06689	0.52193
AI Active	0.29276	0.24452	0.34978	0.31908	0.35965	0.29276	0.36623	0.26206	0.32566	0.26754	0.30811	0.34101	0.29649	0.32851	0.27961	0.30921	0.21491	0.27193
AI Passive	0.37847	0.28704	0.35069	0.21875	0.41204	0.28704	0.29745	0.30556	0.20718	0.41782	0.09838	0.54861	0.17940	0.41551	0.14931	0.40394	0.09243	0.52431
IA Active	0.52604	0.17708	0.78125	0.00000	0.59896	0.18229	0.29688	0.30208	0.35417	0.35417	0.50000	0.36458	0.25000	0.60938	0.09896	0.56250	0.05729	0.64583
IA Passive	0.20417	0.40083	0.22000	0.41167	0.15417	0.46583	0.20833	0.44583	0.13667	0.55833	0.11750	0.55262	0.09917	0.52833	0.10250	0.41250	0.09500	0.45250
Estimate	0.05604	-0.0791	-0.1080	0.11779	-0.1995	0.13623	-0.0209	-0.0595	-0.1024	-0.1510	-0.0618	-0.0197	0.14951	-0.3133	0.06655	-0.1766	0.16203	-0.0622
Std. Error	0.12735	0.13768	0.14159	0.14513	0.12607	0.15096	0.13316	0.14095	0.11236	0.12267	0.10858	0.14571	0.11263	0.14447	0.10492	0.13393	0.09527	0.15261
t value	0.44	-0.574	-0.763	0.812	-1.582	0.902	-0.157	-0.422	-0.912	-1.231	-0.569	-0.135	1.327	-2.169	0.634	-1.319	1.701	-0.408
p value	0.661	0.5661	0.4517	0.4179	0.1176	0.372	0.8756	0.6768	0.365	0.2254	0.5696	0.8925	0.1865	0.03229 *	0.5267	0.1893	0.09158	0.6835

Appendix 5. Placement test of Japanese

SPOT Test is a placement test for Japanese, developed by Tsukuba University (Kobayashi, Niwa & Yamamoto, 1996). It consists of listening to an audio with Japanese speech at normal rate and filling in the blanks. It has shown to be a good predictor of grammatical knowledge in Japanese (Suzuki, 2014). Next we present first the instructions, followed by the test.

SPOT test
Listen to the tape and fill in the blank ( ) with the hiragana symbol.
You will first hear 10 Practice sentences.

①どうぞよろ（　）く。
②ここは静（　）ですね。
③おはよう（　）ざいます。
④わたし（　）たなかです。
⑤ごはんを食（　）ました。
⑥どこから（　）ましたか。
⑦あしたここに来ます（　）。
⑧ぜんぶ（　）いくらですか。
⑨タクシーで行きま（　）ょう。
⑩あたらし（　）車を買いました

The test will now start. The test you will undertake is version (A).When you hear the instrudaction ‘’de wa hajimemasu’’,turn the page and start.

SPOT test

Listen to the tape and fill in the blank (   ) with the hiragana symbol.

(1). そこ（　）何をしてるんですか。
(2). あの人は日本では有名（　）人ですよ。
(3). 今度、映画見（　）行かない？
(4). その中（　）なに入ってんの？
(5). 隣の人（　）教えてもらったんです。
(6). あのグリーン（　）スカート、いいなあ。
(7). 指導教官の先生（　）はもう会いましたか。
(8). 木村先生に会（　）ればいいのですが。
(9). 「TISA」っていうの（　）知っていますか。
(10). あのコーヒー（　）おいしい店、名前なんだっけ？
(11). 明日はちょっと大事（　）の用があって行けないんです。
(12). ほら、あの窓（　）ところにいるでしょう。
(13). 郵便局のところ（　）曲がってください。
(14). あそこに地図がはって（　）りますよ。
(15). 好きな人（　）もいるの？
(16). 肉の色（　）変わったら、火を止めてください。
(17). 彼ったら、偉そう（　）ことばかり言って。
(18). 昨日は一日中そうじ（　）せられて、大変だった。
(19). 約束してたの（　）来なかった。
(20). 旅行の申込書なんですけど、これ（　）いいですか。
(21). 毎日手紙を書く（　）とにしよう。
(22). 来週の会議については、後で（　）連絡します。
(23). 今説明したのが、この茶色（　）見えるところです。
(24). それはそう（　）と思います。
(25). なにやり始める（　　）と思ったら、なあんだ
(26). 会議があったこと、すっかり忘れ（　）た。
(27). 留学生（　）とって筑波は住みにくいところです。
(28). 今日はもうそのぐらい（　）して、早く帰ろう。
(29). 新聞を読んでも、本当の事なかなかわからない（　）けです。
(30). 部屋代は東京（　）ど高くないです。
(31). これからはもっとがんばら（　）きゃ。
(32). 出かけ（　）うとしたら、電話がかかってきた。
(33). 君の（　）いで遅れちゃったよ。
(34). アルバイトっ（　）いえば、この間の話、もう決まった？

(35). 子供（　）なんかわかるわけないだろう。
(36). 必ずしもよくなるとは（　）ぎらない。
(37). うちの母は、もう 60（　）し。
(38). これはうちの問題（　）ありまして、そちらには関係のないことです。
(39). 今の（　）までだいじょうぶでしょう。
(40). あの人、結婚しない（　）じゃないの？
(41). ゴミの問題はひどくなって（　）く一方だ。
(42). 早く行っ（　）って、なんかもらえるわけじゃないし。
(43). 今度のアパート、場所（　）いいんだけどね。
(44). そりゃ、外国人（　）あなたにはいいかもしれないけど。
(45). 今後それをどのように証明できる（　）が、最大のポイントとなります。
(46). 私達もなにかす（　）きだ。
(47). うん、思った（　）りずっと進んでるな。
(48). 就職した（　　）らといって、勉強が終わったというわけじゃないよ。
(49). それだけでは終わりそう（　）ないですね。
(50). 私に言えない（　）うなことでもあるの？
(51). 早く国へ帰りたいなあなん（　）思ったりします。
(52). すみませんが、ちょっと手伝っていただ（　）ませんか。
(53). 作ってはみた（　）のの、あまりいいプログラムじゃなかった。
(54). 私のこと聞いたんでしょう、彼（　）。
(55). あいつ、酒飲んで寝ちゃっ（　）さ。
(56). この調子なら、おれ、どんどん読（　　）ちゃいそう。
(57). 嫌だけど、どうしてもやら（　）るをえないんだ。
(58). 人工が増えるに（　）たがって、住みにくくなってきた。
(59). だから、私はそういう（　）うに思いました。
(60). A：あのう、田中さんという方は？
B：ええと、あそこに立っている人（　）田中さんです。
(61). A：先週山に行ったんですよ。
B：だれ（　）いっしょに出かけたんですか。
(62). A：家事やりますか。
B：せんたく（　）かはしますけど、そうじはしませんね。
(63). A：そのけが、どうしたの。
B：自転車に乗って（　）、ころんじゃった。
(64). A：まいにちひまでひまで。
B：じゃ、明日どこ（　）行かない？。
(65). A：ねえ、この話知ってる？
B：うん、昨日の新聞（　）でてたよ。